Theoretical Population Biology 78 (2010) 123–138

Contents lists available at ScienceDirect

Theoretical Population Biology

journal homepage: www.elsevier.com/locate/tpb

Evolution and polymorphism in the multilocus Levene model with no or weak epistasis

Reinhard Bürger *

Department of Mathematics, University of Vienna, Austria

ARTICLE INFO

Article history:
Received 29 January 2010
Available online 16 June 2010

Keywords:
Selection
Migration
Recombination
Linkage disequilibrium
Population subdivision
Geographical structure
Quantitative traits
Genotype–environment interaction

ABSTRACT

Evolution and the maintenance of polymorphism under the multilocus Levene model with soft selection are studied. The number of loci and alleles, the number of demes, the linkage map, and the degree of dominance are arbitrary, but epistasis is absent or weak. We prove that, without epistasis and under mild, generic conditions, every trajectory converges to a stationary point in linkage equilibrium. Consequently, the equilibrium and stability structure can be determined by investigating the much simpler gene-frequency dynamics on the linkage-equilibrium manifold. For a haploid species an analogous result is shown. For weak epistasis, global convergence to quasi-linkage equilibrium is established. As an application, the maintenance of multilocus polymorphism is explored if the degree of dominance is intermediate at every locus and epistasis is absent or weak. If there are at least two demes, then arbitrarily many multiallelic loci can be maintained polymorphic at a globally asymptotically stable equilibrium. Because this holds for an open set of parameters, such equilibria are structurally stable. If the degree of dominance is not only intermediate but also deme independent, and loci are diallelic, an open set of parameters yielding an internal equilibrium exists only if the number of loci is strictly less than the number of demes. Otherwise, a fully polymorphic equilibrium exists only nongenerically, and if it exists, it consists of a manifold of equilibria. Its dimension is determined. In the absence of genotype-by-environment interaction, however, a manifold of equilibria occurs for an open set of parameters. In this case, the equilibrium structure is not robust to small deviations from no genotype-by-environment interaction. In a quantitative-genetic setting, the assumptions of no epistasis and intermediate dominance are equivalent to assuming that in every deme directional selection acts on a trait that is determined additively, i.e., by nonepistatic loci with dominance. Some of our results are exemplified in this quantitative-genetic context.

1. Introduction

To achieve a proper understanding of the evolutionary dynamics of phenotypic traits, it is essential to study the effects of selection on multiple linked or unlinked loci. Because many species are subdivided into colonies, or demes, and selection varies geographically, the consequences of migration and spatially varying selection need to be taken into account. Each of these aspects has been studied extensively but mostly separately. Multilocus selection and the maintenance of polygenic variation in a panmictic population inhabiting a constant, homogeneous environment have been prime topics of research during the past decades; for reviews or recent treatments of general models, see Karlin (1978), Turelli and Barton (1990), Lyubich (1992), Nagylaki (1992), Zhivotovsky and Gavrilets (1992), Christiansen (1999), Bürger (2000), Kirkpatrick et al. (2002), and Ewens (2004).

Spatially varying selection in subdivided populations was intensively investigated as well, mainly for the single-locus case. A particularly prominent role has been played by the Levene (1953) model, which assumes a finite number of demes, selection within demes, and individuals dispersing independently of their deme of origin. As a consequence, there is no population structure despite geographically variable selection. This property makes it more amenable to mathematical analysis than general migration–selection models. A good overview of the literature can be acquired from the articles of Karlin (1977, 1982) and Nagylaki and Lou (2008), and the pertinent chapters in the books of Nagylaki (1992) and Christiansen (1999).

Work on multilocus selection in subdivided populations is relatively scarce. There is early work by Li and Nei (1974), who showed that even in the absence of epistasis and dominance, migration–selection balance in two demes can maintain linkage disequilibrium (see also Christiansen and Feldman, 1975). Zhivotovsky et al. (1996) used a multilocus Levene model to study the evolution of phenotypic plasticity. Wiehe and Slatkin (1998) investigated a haploid Levene model in which linkage disequilibrium is caused

* Corresponding address: Institut für Mathematik, Universität Wien, Nordbergstrasse 15, A-1090 Wien, Austria.
E-mail address: reinhard.buerger@univie.ac.at.

0040-5809
doi:10.1016/j.tpb.2010.06.002

by epistasis. Christiansen (1999) derived sufficient conditions for the protection of gametes in a multilocus context. More recently, Spichtig and Kawecki (2004) and van Doorn and Dieckmann (2006) performed numerical studies on the maintenance of multilocus polymorphism in two demes for arbitrary migration and for the Levene model, respectively. Roze and Rousset (2008) derived recursions for the allele frequencies and for various types of genetic associations in a multilocus infinite-island model. Barton (2010) explored certain aspects of speciation using a generalized, haploid multilocus Levene model that admits habitat preferences.

The notorious complexity of the evolutionary dynamics of multilocus systems as well as the richness of evolutionary phenomena in subdivided populations leave little hope for a general theory combining both aspects. However, considerable progress has been made recently for important limiting or special cases. These include weak migration and weak selection (Bürger, 2009a,b) and the Levene model without epistasis (Nagylaki, 2009b; Bürger, 2009c). If either migration or selection is weak, the evolutionary dynamics are perturbations of relatively simple limiting dynamics which are amenable to mathematical analysis. In both cases, global convergence of trajectories to equilibria in quasi-linkage equilibrium could be proved under natural, quite general assumptions (Bürger, 2009a). If migration is weak, this conclusion requires the assumption of weak epistasis; if selection is weak, then equilibrium states are also spatially quasi-homogeneous.

These results were applied in Bürger (2009b) to study the maintenance of multilocus polymorphism if epistasis is weak or absent and dominance is intermediate. In a panmictic population, polymorphism is impossible under such conditions. For strong migration, however, arbitrarily many recombining loci can be maintained polymorphic if there are at least two demes, and this holds for an open set of parameters. By contrast, for weak migration, the maximum number of loci that can be maintained polymorphic on an open set of parameters equals the number of demes.

Nagylaki (2009b) investigated evolution under the multiallelic multilocus Levene model without epistasis. He demonstrated that geometric-mean fitness, $\tilde{w}(\rho)$, depends only on the vector ρ of gene frequencies and is monotone increasing except at equilibria. Therefore, $\rho(t)$ converges generically, i.e., for almost all parameters and initial data, to a gene-frequency equilibrium that is a local maximum of $\tilde{w}(\rho)$. In addition, Nagylaki proved global convergence to linkage equilibrium if there are either only two loci or there are multiple loci without dominance. He conjectured that the set of gene-frequency equilibria that are in linkage equilibrium is globally attracting, hence that global convergence to linkage equilibrium occurs always. For deme-independent degree of intermediate dominance (DIDID) he showed that, generically, at most $\Gamma - 1$ diallelic loci can segregate at equilibrium, where Γ denotes the number of demes.

In Bürger (2009c), the equilibrium and stability structure of the diallelic two-locus Levene model with two demes was derived in considerable generality. Epistasis was ignored but dominance admitted. Absence of genotype-by-environment ($G\times E$) interaction was shown to lead to nongeneric, and nonrobust, properties.

In this paper, we shall prove Nagylaki's conjecture for multiple multiallelic loci under the generic assumption that gene-frequency equilibria are isolated. As a consequence, evolution in the Levene model without epistasis can be fully understood by studying the much simpler gene-frequency dynamics on the linkage-equilibrium manifold for which geometric-mean fitness is monotone increasing along nonconstant solutions. More generally, we establish convergence of trajectories to a stationary point in quasi-linkage equilibrium if epistasis is sufficiently weak. Analogous, and even stronger, results hold if selection acts on haploids. We apply these results to investigate the maintenance of multilocus polymorphism in the Levene model for diallelic nonepistatic loci and, especially, for a quantitative trait that is under linear selection in every deme.

In Section 2, we formulate the multilocus Levene model and summarize the results on the gene-frequency dynamics that are needed subsequently. Section 3 is devoted to the study of convergence to linkage equilibrium in the absence of epistasis. The main result is Theorem 3.1, which states global convergence to stationary points in linkage equilibrium under mild, generic conditions. The proof is based on the simple observation that, at gene-frequency equilibrium, the dynamics reduces to a panmictic dynamics without epistasis. The theorem follows by employing the results of Nagylaki (2009b) on convergence of gene frequencies in the Levene model and those of Kun and Lyubich (1979, 1980) on convergence in the panmictic case without epistasis. Corollary 3.3 is most useful for applications because it formulates several of the conclusions that can be deduced for the full dynamics by analyzing the much simpler gene-frequency dynamics at linkage equilibrium. In Section 3.2, geometric convergence to a unique, globally asymptotically stable equilibrium is established under the assumption of DIDID (Theorem 3.5). Among others, this includes absence of $G \times E$ interaction as a special case. In Section 4, the haploid Levene model is studied and geometric convergence to a unique, globally asymptotically stable equilibrium is proved.

The maintenance of multilocus polymorphism is investigated in Section 5. Result 5.1 is a slight extension of Theorem 2.2 in Bürger (2009b) which shows that in the Levene model an arbitrary number of loci can be polymorphic at a globally asymptotically stable equilibrium, and this holds for an open set of parameters. The main results are Theorems 5.3 and 5.5. They apply if the degree of dominance is intermediate and deme independent and show that this assumption considerably restricts the possibility of multilocus polymorphism relative to intermediate dominance that varies among demes. The first theorem establishes that Nagylaki's (2009b) generic upper bound, $\Gamma - 1$, for the number of segregating loci at equilibrium is assumed on an open set of parameters. The second theorem shows that in the nongeneric case, in which an internal equilibrium with more than $\Gamma - 1$ polymorphic loci exists, there is a manifold of equilibria with generic dimension $L - \Gamma + 1$, where L is the number of loci. This highly degenerate case occurs for instance in the absence of $G \times E$ interaction.

Section 6 applies these results to a quantitative-genetic model with linear directional selection in each deme. Corollary 6.3, which summarizes the main results, should serve as a warning when studying models under highly specialized assumptions. It shows that natural assumptions, such as absence of genotype-by-environment interaction, may lead to nongeneric model behavior. Indeed, the analysis of such a simple degenerate model (two diallelic loci, two demes, no dominance and linear fitnesses; see Bürger, 2009c) and the ensuing discussions with Thomas Nagylaki initiated the recent series of papers on multilocus migration–selection models by Nagylaki and the author.

In Section 7, weak epistasis is studied. In Section 7.1, we establish a perturbation result that yields global convergence of trajectories to quasi-linkage equilibrium. In Section 7.2, some of the results of Section 5 on the maintenance of polymorphism are extended to weak epistasis. In Section 8, we recapitulate and discuss our main findings, and mention some open problems.

2. The multilocus Levene model

We briefly introduce the multilocus Levene model and summarize some basic results from Nagylaki (2009b), hereafter abbreviated as N09b, that will be needed. There the model is developed in detail.

We suppose that there are $L \geq 2$ diploid loci with $I_n \geq 2$ alleles $A_{i_n}^{(n)}$ ($i_n \in \{1, \ldots I_n\}$) at locus n. We designate the set of all loci by

Table 1
Glossary of symbols. For both the Roman and Greek alphabets, uppercase letters precede lowercase ones. For each uppercase or lowercase letter, listing is in order of appearance of the definition in the text. The references are to the equation closest to the definition of each symbol. Thus, (2.1), (2.1)+, (2.1)− refers to (2.1), the text below (2.1), the text above (2.1), respectively. Symbols that occur only in the Appendix are not listed.

Symbol	Reference	Definition
$A_{i_n}^{(n)}$	(2.1)-	Allele at locus n
$\hat{A}$	(3.4)+	Region of attraction of $\hat{\rho}$
b	(5.17)	$(b_1, \ldots, b_\Gamma)^T$
b_α	(5.17)+	$\hat{\bar{w}}_\alpha - \sum_n u_{22,\alpha}^{(n)}$
c_α	(2.1)-	Proportion of zygotes in deme α
D	(3.4)-	Vector of all linkage disequilibria in an L-locus system
$D_{i,\alpha}$	(3.4)-	Linkage disequilibrium in gamete i in deme α
$\hat{D}$	(3.4)+	D evaluated at gene-frequency equilibrium
F	(2.14)	$\ln \tilde{w}$
f	(5.14)-	$(f^{(1)}, \ldots, f^{(L)})^T$
$f^{(n)}$	(5.14)	$2\vartheta^{(n)}p^{(n)} + (1 - 2\vartheta^{(n)})(p^{(n)})^2$
G	(2.1)-	Set of all demes
$G^{(n)}$	(2.17)-	Covariance matrix for locus n
h	(2.1)-	Locus index
I_n	(2.1)-	Number of alleles at locus n
I	(6.8)-	Number of alleles at locus n if independent of n
i_n	(2.1)-	Allelic index at locus n
i	(2.1)-	Gamete $(i_1, \ldots, i_L)$
J	(2.3)-	Number of gametes
J_n	(7.3)-	Set of alleles present at equilibrium at locus n
j	(2.1)-	Index for gamete
j_n	(2.4)-	Allelic index at locus n
k	(2.1)-	Index for gamete
L	(2.1)-	Number of loci
L	(2.1)-	Set of all loci
l	(2.1)-	Locus index
m_α	(6.7)+	Backward migration rate
n	(2.1)-	Locus index
O	(7.3)+	Order symbol
p_i	(2.1)-	Frequency of gamete i
$p_{i_n}^{(n)}$	(2.1)-	Frequency of allele $A_{i_n}^{(n)}$
p	(2.3)-	Vector of gamete frequencies
$p_{j_h i_n}^{(hn)}$	(2.10)	Frequency of gamete $A_{j_h}^{(h)}A_{i_n}^{(n)}$
$p^{(n)}$	(2.17)-	$(p_1^{(n)}, \ldots, p_{I_n}^{(n)})^T$ (only Section 2)
$p^{(n)}$	(5.1)-	Frequency of allele $A_1^{(n)}$ in the diallelic case
$\hat{p}$	(3.4)+	p at gamete-frequency equilibrium
$R_{i,jk}$	(2.1)+	Probability that gamete i is produced by haplotypes j, k
$\mathbb{R}^J$	(2.3)-	J-dimensional Euclidean space
$r_{\min}$	(3.8)+	Smallest two-locus recombination frequency
s_α	(6.2)	Directional selection coefficient in deme α
$s_{ij,\alpha}$	(7.1)	Epistasis coefficients
t	(2.10)+	Time in generations
U	(5.9)+	Open set of parameters
$u_{i_n j_n,\alpha}^{(n)}$	(2.3)+	Fitness contribution of $A_{i_n}^{(n)}A_{j_n}^{(n)}$ in deme α
$\bar{u}_\alpha^{(n)}$	(2.6)	Mean fitness contribution of locus n in deme α
$u_{i_n,\alpha}^{(n)}$	(2.7)	Marginal fitness contribution of $A_{i_n}^{(n)}$ in deme α
$u_{i_n,\alpha}^{(n)}$	(4.2a)	Fitness contribution of $A_{i_n}^{(n)}$ in deme α (only Section 4)
V	(5.10)-	$L \times \Gamma$ matrix with entries $V_{n\alpha}$
$V_{n\alpha}$	(5.10)-	Entry of matrix V
W_{jk}	(3.1)	Averaged fitness of jk at gene-frequency equilibrium
$\overline{W}_\alpha(z)$	(5.15)	Mean fitness in deme α as a function of z
$W_\alpha(X)$	(6.2)	Fitness of individuals with trait value X in deme α
W	(7.3)+	Open set of parameters
$w_{ij,\alpha}$	(2.1)-	Fitness of genotype ij in deme α
$w_{i,\alpha}$	(2.1)	Marginal fitness of gamete i in deme α
$w_{i,\alpha}$	(4.1)-	Fitness of gamete i in deme α (only Section 4)
$\overline{w}_\alpha$	(2.1)	Mean fitness in deme α
$\tilde{w}$	(2.13)	Geometric-mean fitness
$\hat{\bar{w}}_\alpha$	(3.1)-	$\overline{w}_\alpha(\hat{\rho})$
X	(6.1)	Value of quantitative trait
$X_{ij,\alpha}$	(6.1)	$\sum_n \gamma_{i_n j_n,\alpha}^{(n)}$
$\hat{x}$	(5.11)+	$(c_1/\hat{\bar{w}}_1, \ldots, c_\Gamma/\hat{\bar{w}}_\Gamma)^T$
$z^{(n)}$	(5.14)	$f^{(n)}(p^{(n)})$
z	(5.14)+	$(z^{(1)}, \ldots, z^{(L)})^T$
α	(2.1)-	Deme index
β	(2.1)-	Deme index
Γ	(2.1)-	Number of demes
$\gamma_{i_n j_n,\alpha}^{(n)}$	(6.1)+	Contribution of $A_{i_n}^{(n)}A_{j_n}^{(n)}$ to trait value

(*continued on next page*)

Table 1 (*continued*)

Symbol	Reference	Definition
$\gamma^{(n)}_{i_n j_n}$	(6.6)	Deme-independent contribution of $A^{(n)}_{i_n}A^{(n)}_{j_n}$ to trait value
γ_α	(6.6)	Genotype-independent contribution to trait value
Δ_J	(2.3)-	Simplex in $\mathbb{R}^J$
Δ	(2.15)-	Difference operator
δ_{ij}	(2.17)-	Kronecker delta
$\vartheta^{(n)}_{i_n j_n}$	(3.6)	Degree of intermediate dominance
$\vartheta^{(n)}_{i_n j_n}$	(6.5)	Degree of intermediate dominance at trait level
$\vartheta^{(n)}$	(5.7)	Degree of intermediate dominance in the diallelic case
ϑ	(5.9)+	$(\vartheta^{(1)},\dots,\vartheta^{(L)})^T$
ϑ	(6.15)-	ϑ_{12}
$\vartheta^{(n)}_{i_n}$	(6.8)	Degree of intermediate dominance of $A^{(n)}_{i_n}$ at the trait level
ϵ	(7.1a)-	Small positive parameter
Λ	(2.11)	$\{p \in \Delta_J : \rho' = \rho\}$
$\hat{\Lambda}$	(2.18)	$\{p \in \Delta_J : \rho = \hat{\rho}\}$
Ξ	(2.12)	Linkage-equilibrium manifold
$\xi^{(n)}_{i_n j_n}$	(2.8a)	Averaged relative fitness of $A^{(n)}_{i_n}A^{(n)}_{j_n}$
$\hat{\xi}^{(n)}_{i_n j_n}$	(2.18)+	$\xi^{(n)}_{i_n j_n}$ at gene-frequency equilibrium
$\xi^{(n)}_{i_n}$	(2.8b)	Averaged relative fitness of $A^{(n)}_{i_n}$
$\hat{\xi}^{(n)}_{i_n}$	(2.18)+	$\xi^{(n)}_{i_n}$ at gene-frequency equilibrium
$\bar{\xi}^{(n)}$	(2.8c)	Averaged mean fitness at locus n
$\hat{\bar{\xi}}^{(n)}$	(2.18)+	$\bar{\xi}^{(n)}$ at gene-frequency equilibrium
$\xi^{(n)}_{i_n}$	(4.3)	Averaged relative fitness of $A^{(n)}_{i_n}$ (only Section 4)
ρ	(2.3)-	Vector of allele frequencies
ρ	(5.1)	Vector of allele frequencies (diallelic case)
$\hat{\rho}$	(2.18)-	ρ at gene-frequency equilibrium
Σ_0	(7.1b)+	Set of equilibria without epistasis
$\Sigma_{w(\epsilon)}$	(7.1b)+	Set of equilibria with epistasis
$\phi^{(n)}_{i_n}$	(6.11)	Allele specific factor of fitness excess
Ω	(2.3)	Gene-frequency space
Ω	(5.1)	Gene-frequency space (diallelic case)
$'$	(2.2)	Value of quantity in next generation
$\subseteq$	(2.3)-	Subset
^{T}	(2.3)-	Transposition of a vector
int	(5.9)+	Topological interior of a set
rank	(5.12)	Rank of a matrix
dim	(5.11)+	Dimension of a linear subspace
ker	(5.11)+	Kernel of a linear map or matrix
$\nabla^{(n)}$	(2.17)+	Gradient operator

$\mathsf{L} = \{1,\dots,L\}$ and the set of all demes by $\mathsf{G} = \{1,\dots,\Gamma\}$, where $\Gamma \geq 2$ is assumed. The relative size of deme α is denoted by $c_\alpha > 0$, hence $\sum_\alpha c_\alpha = 1$. (Whenever no summation range is indicated, it is assumed to be over all admissible values; here, $\alpha \in \mathsf{G}$.) We shall consistently use the letters $h, l, n \in \mathsf{L}$ for loci, i, j, k for gametes, and $\alpha, \beta \in \mathsf{G}$ for demes (see Table 1 for a glossary of symbols). The linkage map is arbitrary, except for the assumption that all recombination probabilities are positive.

Throughout, we assume the Levene model with soft selection, which means that population regulation by selection occurs within demes. This assumption induces frequency-dependent selection. Because in the Levene model migration rates are independent of the deme of origin, there is no population structure, and gamete and gene frequencies before selection are deme independent (Levene, 1953; Nagylaki, 1992). We denote the frequency of gamete i, which carries allele $A^{(n)}_{i_n}$ at locus n, by p_i, and the frequency of allele $A^{(n)}_{i_n}$ by $p^{(n)}_{i_n}$. Let $w_{ij,\alpha}$ be the fitness of the diploid genotype ij in deme α. Then the marginal fitness of gamete i and the mean fitness of the population in deme α are

$$w_{i,\alpha} = \sum_j p_j w_{ij,\alpha} \quad \text{and} \quad \overline{w}_\alpha = \sum_{i,j} p_i p_j w_{ij,\alpha}, \tag{2.1}$$

respectively. Further, let $R_{i,jk}$ denote the probability that a parent of genotype jk produces a gamete i during meiosis. Because there is soft selection, adult dispersal, and random mating within demes, the gamete frequencies evolve according to (N09b, Eq. (2.42))

$$p'_i = \sum_{j,k,\alpha} R_{i,jk} p_j p_k c_\alpha w_{jk,\alpha} / \overline{w}_\alpha, \tag{2.2}$$

where the prime, $'$, signifies the next generation. We note that these recursions are obtained whether dispersion precedes recombination or not (Bürger, 2009a). Moreover, they admit the classical interpretation that intrademic selection is followed by random mating in the entire population (cf. Nagylaki and Lou, 2008).

The state space is the simplex $\Delta_J \subseteq \mathbb{R}^J$ of probability vectors of length J, where $J = \prod_n I_n$ is the number of gametes. We write $p = (p_1,\dots,p_J)^T \in \Delta_J$ for the vector of gametic frequencies. The vector consisting of all gene frequencies $p^{(n)}_{i_n}$ (for every n and every i_n) is denoted by $\rho \in \Omega$, where

$$\Omega = \prod_n \Delta_{I_n} \tag{2.3}$$

is the space of gene frequencies, or the gene-frequency space, for short.

For the rest of this section, we assume *absence of epistasis*. Then we can assign fitness contributions to single-locus genotypes. We denote the contributions at locus n in deme α by $u^{(n)}_{i_n j_n,\alpha}$, where i_n and j_n refer to the alleles carried by the genotype ij at locus n, and assume $u^{(n)}_{i_n j_n,\alpha} = u^{(n)}_{j_n i_n,\alpha} \geq 0$. Thus, we posit that the fitness of genotype ij is given by

$$w_{ij,\alpha} = \sum_n u^{(n)}_{i_n j_n,\alpha}. \tag{2.4}$$

An easy calculation shows that the mean fitness in deme α becomes

$$\overline{w}_\alpha = \sum_n \bar{u}^{(n)}_\alpha, \tag{2.5}$$

where

$$\bar{u}_\alpha^{(n)} = \sum_{i_n,j_n} u_{i_nj_n,\alpha}^{(n)} p_{i_n}^{(n)} p_{j_n}^{(n)} = \sum_{i_n} u_{i_n,\alpha}^{(n)} p_{i_n}^{(n)}, \tag{2.6}$$

and

$$u_{i_n,\alpha}^{(n)} = \sum_{j_n} u_{i_nj_n,\alpha}^{(n)} p_{j_n}^{(n)} \tag{2.7}$$

is the fitness contribution of allele $A_{i_n}^{(n)}$ in deme α. Importantly, $\overline{w}_\alpha = \overline{w}_\alpha(\rho)$ depends only on the vector ρ of gene frequencies, but not on the vector p of gamete frequencies.

We introduce the compact notation

$$\xi_{i_nj_n}^{(n)} = \sum_\alpha \frac{c_\alpha}{\overline{w}_\alpha} u_{i_nj_n,\alpha}^{(n)}, \tag{2.8a}$$

$$\xi_{i_n}^{(n)} = \sum_{j_n} \xi_{i_nj_n}^{(n)} p_{j_n}^{(n)} = \sum_\alpha \frac{c_\alpha}{\overline{w}_\alpha} u_{i_n,\alpha}^{(n)}, \tag{2.8b}$$

$$\bar{\xi}^{(n)} = \sum_{i_n} \xi_{i_n}^{(n)} p_{i_n}^{(n)} = \sum_\alpha \frac{c_\alpha}{\overline{w}_\alpha} \bar{u}_\alpha^{(n)}, \tag{2.8c}$$

and note that $\xi_{i_nj_n}^{(n)}$, $\xi_{i_n}^{(n)}$, and $\bar{\xi}^{(n)}$ are nonnegative functions of ρ, and

$$\sum_n \bar{\xi}^{(n)} = 1. \tag{2.9}$$

With this notation, the dynamics of gene frequencies (Eq. (2.48) in N09b) can be written as

$$p_{i_n}^{(n)\prime} = p_{i_n}^{(n)} \xi_{i_n}^{(n)} + \sum_{h:h\neq n} \sum_{j_h} p_{j_hi_n}^{(hn)} \xi_{j_h}^{(h)}, \tag{2.10}$$

where $p_{j_hi_n}^{(hn)}$ is the frequency of $A_{j_h}^{(h)} A_{i_n}^{(n)}$. Clearly, every solution $p(t)$ of (2.2) generates a solution $\rho(t)$ of (2.10), but ρ' cannot be inferred from (2.10) if only ρ, and not p, is known.

We define

$$\Lambda = \{p \in \Delta_J : \rho' = \rho\} \tag{2.11}$$

as the set of gametic frequencies at gene-frequency equilibrium, or the set of gene-frequency equilibria for short, and

$$\Xi = \{p \in \Delta_J : p_i = p_{i_1}^{(1)} \cdot \ldots \cdot p_{i_L}^{(L)} \text{ for every gamete } i\} \tag{2.12}$$

as the linkage-equilibrium manifold. Further, let

$$\tilde{w}(\rho) = \prod_\alpha [\bar{w}_\alpha(\rho)]^{c_\alpha} \tag{2.13}$$

denote the geometric-mean fitness and define

$$F(\rho) = \ln \tilde{w}(\rho) = \sum_\alpha c_\alpha \ln \overline{w}_\alpha(\rho). \tag{2.14}$$

The following result will play a fundamental role. It depends crucially on the fact that $\tilde{w}$ and F are functions only of ρ rather than of p.

Result 2.1 (*N09b*)**.** *In the absence of epistasis, the dynamics (2.2) of gamete frequencies has the following properties:*

(a) *The single-generation change $\Delta\tilde{w}(\rho) = \tilde{w}(\rho') - \tilde{w}(\rho)$ satisfies $\Delta\tilde{w}(\rho) \geq 0$, and equality holds only at gene-frequency equilibrium. The function $F(\rho)$ shares these properties.*

(b) *The internal gene-frequency equilibria are the stationary points of $\tilde{w}$. These are obtained as the critical points of F and satisfy*

$$\xi_{i_n}^{(n)} = \bar{\xi}^{(n)} \quad \textit{for every } n \textit{ and every } i_n. \tag{2.15}$$

Statement (a) about $\Delta\tilde{w}$ is Theorem 3.1 in N09b. The statement about ΔF is a trivial consequence. Statement (b) is Theorem 3.3 in N09b.

We call a property generic if it holds in an open dense set of full measure.

Remark 2.2. Simple important consequences of statement (a) are (Remark 3.2 in N09b):

(i) The set $\Lambda \subseteq \Delta_J$ of gene-frequency equilibria is globally attracting, i.e., $p(t) \to \Lambda$ as $t \to \infty$.

(ii) If every equilibrium point $\hat{\rho}$ is isolated in the gene-frequency space Ω, as is generic, every $\rho(t)$ converges as $t \to \infty$ to some $\hat{\rho} \in \Omega$.

(iii) For generic initial conditions $\rho(t)$ converges to a local maximum of $\tilde{w}$.

If there is linkage equilibrium, i.e., on Ξ, the recursions (2.10) for the gene frequencies simplify to

$$\Delta p_{i_n}^{(n)} = p_{i_n}^{(n)} \sum_\alpha \frac{c_\alpha}{\overline{w}_\alpha} (u_{i_n}^{(n)} - \bar{u}_\alpha^{(n)}) = p_{i_n}^{(n)} (\xi_{i_n}^{(n)} - \bar{\xi}^{(n)}). \tag{2.16}$$

In fact, (2.16) can be written as a generalized gradient system, similar to the single-locus selection dynamics in a panmictic population (Bürger, 2000, p. 42). To see this, note that

$$\frac{\partial F}{\partial p_{i_n}^{(n)}} = 2 \sum_\alpha \frac{c_\alpha}{\overline{w}_\alpha} u_{i_n}^{(n)} = 2\xi_{i_n}^{(n)},$$

and define $p^{(n)} = (p_1^{(n)}, \ldots, p_{I_n}^{(n)})^T$ and $G^{(n)}$ as the $I_n \times I_n$ covariance matrix with entries $\frac{1}{2} p_{i_n}^{(n)} (\delta_{i_nj_n} - p_{j_n}^{(n)})$. Then, for every locus n, we obtain

$$\Delta p^{(n)} = G^{(n)} \nabla^{(n)} F, \tag{2.17}$$

where $\nabla^{(n)} F = (\partial F/\partial p_1^{(n)}, \ldots, \partial F/\partial p_{I_n}^{(n)})^T$. Although, (2.16) and (2.17) are closed systems on Ω, it should be noted that Ξ is not invariant under the full dynamics (2.2), as this is not even the case in a panmictic population.

Remark 2.3. $F(\rho)$ and $\tilde{w}(\rho)$ are Lyapunov functions for (2.16), and for (2.17), and the asymptotically stable equilibria of (2.16) are the local maxima of $F(\rho)$. The internal equilibria of (2.16) are exactly the internal critical points of $F(\rho)$, or of $\tilde{w}(\rho)$, on Δ_J, and this holds for every lower-dimensional subsystem. They are precisely the solutions of (2.15). Comparison with Result 2.1 yields that $\hat{\rho}$ is an (asymptotically stable) equilibrium of (2.16) if and only if it gives rise to an (asymptotically stable) gene-frequency equilibrium of (2.2).

We designate equilibrium points of (2.16) by $\hat{\rho} \in \Omega$. Every $\hat{\rho}$ gives rise to the set

$$\hat{\Lambda} = \{p \in \Delta_J : \rho = \hat{\rho}\} \subseteq \Lambda \tag{2.18}$$

of gene-frequency equilibria of (2.2). We write $\hat{\xi}_{i_nj_n}^{(n)}$, $\hat{\xi}_{i_n}^{(n)}$, and $\hat{\bar{\xi}}^{(n)}$ if $\xi_{i_nj_n}^{(n)}$, $\xi_{i_n}^{(n)}$, and $\bar{\xi}^{(n)}$, respectively, are evaluated at $\hat{\rho}$.

We will need the following two assumptions on the gene-frequency equilibria:

The equilibrium points $\hat{\rho}$ of (2.16) are isolated in Ω. (2.19)

$\hat{\xi}_{i_nj_n}^{(n)} > 0$ for every n and every i_n and j_n. (2.20)

Clearly, (2.20) is satisfied if no single-locus genotype is lethal everywhere. A simple consequence is $\hat{\bar{\xi}}^{(n)} > 0$. We recall that we also assume

All recombination probabilities are positive. (2.21)

These three conditions hold generically.

3. Convergence to linkage equilibrium

In the absence of epistasis, generic global convergence to $\Xi \cap \Lambda$, hence to a stationary point in linkage equilibrium, was proved in N09b if there are two multiallelic loci (Theorem 4.6) or if the number of loci is arbitrary and there is no dominance (Theorem 4.14). Theorem 4.13 in N09b states (local) asymptotic stability of $\Xi \cap \Lambda$ for multiple multiallelic loci. However, as Professor Nagylaki informed me, the induction proof of this theorem has a

gap that may require extensive calculations to fill. (The gap does not affect the validity of the proof of his Theorem 4.14 for no dominance.) Motivated by this failure, I found a simple alternative method that is more general and yields global convergence to linkage equilibrium for the general multiallelic multilocus model without epistasis. It is based on a simple observation and employs the results in N09b on the gene-frequency dynamics (summarized in Result 2.1) as well as the result of Kun and Lyubich (1979, 1980) on convergence to equilibrium in the panmictic multilocus model without epistasis. Throughout this section, we assume absence of epistasis.

3.1. The general case

The crucial observation is the following. Because there is no epistasis, $\overline{w}_\alpha$ is a function of the gene frequencies only. For a given gene-frequency equilibrium $\hat{\rho} \in \Omega$, we write $\hat{\bar{w}}_\alpha = \overline{w}_\alpha(\hat{\rho})$ and obtain from (2.4) and (2.8a):

$$\sum_\alpha c_\alpha \frac{w_{jk,\alpha}}{\hat{\bar{w}}_\alpha} = \sum_\alpha \frac{c_\alpha}{\hat{\bar{w}}_\alpha} \sum_n u^{(n)}_{j_n k_n,\alpha} = \sum_n \hat{\xi}^{(n)}_{j_n k_n}. \tag{3.1}$$

On $\hat{\Lambda}$, every $\hat{\xi}^{(n)}_{j_n k_n}$ is constant because every $\hat{\bar{w}}_\alpha$ is. Therefore, on $\hat{\Lambda}$, the dynamics (2.2) of gamete frequencies becomes

$$p_i' = \sum_{j,k} R_{i,jk} p_j p_k W_{jk}, \tag{3.2}$$

where

$$W_{jk} = \sum_n \hat{\xi}^{(n)}_{j_n k_n} \tag{3.3}$$

is constant. Hence, (3.2) is equivalent to a panmictic multilocus selection dynamics without epistasis. It describes the (panmictic) evolution of linkage disequilibria under selection. Global convergence of trajectories to an equilibrium point in linkage equilibrium now follows from Theorems 1 and 2 in Kun and Lyubich (1979); for a detailed treatment in English, see Section 9.6 in Lyubich (1992). (Application of their result requires the assumptions (2.20) and (2.21).) Because the equilibria of the gene-frequency dynamics are isolated, every $\rho(t)$ converges as $t \to \infty$ to some $\hat{\rho} \in \Omega$. Hence, for every solution $p(t)$ of (2.2), there exists a $\hat{\rho}$ such that the limit points of $p(t)$ are contained in $\hat{\Lambda}$ (Remark 2.2). Now LaSalle's invariance principle (LaSalle, 1977, p. 10) yields the following theorem:

Theorem 3.1. *If there is no epistasis and the assumptions (2.19), (2.20), (2.21) are satisfied, then every trajectory $p(t)$ of the dynamics (2.2) of gamete frequencies converges to an equilibrium point that is in linkage equilibrium. Therefore, $\Lambda \cap \Xi$ is the global attractor for solutions of (2.2).*

Remark 3.2. Theorem 3.1 holds more generally if every trajectory $\rho(t)$ converges to an equilibrium point $\hat{\rho}$. However, if a manifold of equilibria exists in Ω, convergence has not been proved. A proof would require a generalization of the argument that establishes Theorem 9.6.3 in Lyubich (1992), which seems very challenging.

Because we have shown convergence to linkage equilibrium under generic conditions, the full dynamics of gamete frequencies as well as the equilibrium and stability structure can be determined by studying the much simpler gene-frequency dynamics (2.16). To formulate the result precisely, we denote the vector of all linkage disequilibria in the L-locus system by $D = (D_{i,\alpha})$, where $D_{i,\alpha}$ is the linkage disequilibrium defined in Eq. (4.47) of N09b. Then we can write

$$p = (\rho, D). \tag{3.4}$$

Corollary 3.3. *Under the assumptions of Theorem 3.1 the following hold:*

(a) $\hat{p} = (\hat{\rho}, \hat{D})$ *is an equilibrium of (2.2) if and only if $\hat{\rho}$ is an equilibrium of (2.16) and $\hat{D} = 0$, i.e., if and only if $\hat{p} = \hat{\Lambda} \cap \Xi$.*

(b) *$\hat{p}$ is an asymptotically stable equilibrium of (2.2) if and only if $\hat{\rho}$ is an asymptotically stable equilibrium of (2.16). This equivalence also holds for globally asymptotically stable equilibrium points.*

(c) *If $F(\rho)$ is concave and $\hat{\rho}$ is an isolated equilibrium of (2.16) that is either stable or internal, then $(\hat{\rho}, 0)$ is the globally asymptotically stable equilibrium of (2.2).*

Proof. (a) This is an immediate consequence of Theorem 3.1 and Remark 2.3.

(b) For given $\hat{\rho}$, $(\hat{\rho}, 0)$ is an asymptotically stable equilibrium of the dynamics (3.2) on $\hat{\Lambda}$ because Kun and Lyubich (1979) proved convergence to $D = 0$ with the help of a Lyapunov function (see, e.g., Theorem 5.7 in LaSalle, 1977). In view of Remark 2.3, the statement about asymptotic stability is then an immediate consequence of Theorem 3.1 because asymptotic stability of $\hat{\rho}$ implies isolation. If the region of attraction of $\hat{\rho}$ with respect to (2.16) is $\hat{A}$, the region of attraction of $\hat{p}$ with respect to (2.2) is $\{p \in \Delta_J : \rho \in \hat{A}\}$. This yields the statement on global asymptotic stability.

(c) Corollary 3.8 in N09b and the subsequent remark establish uniqueness of a stable gene-frequency equilibrium and its global stability for the gene-frequency dynamics. Convergence to linkage equilibrium follows from Theorem 3.1. □

Corollary 3.3(c) applies in particular if there is no dominance (cf. Theorem 4.14 in N09b). For diallelic loci, it applies whenever fitness contributions at every locus are sublinear, i.e., if for every n and every α,

$$u^{(n)}_{12,\alpha} \geq \frac{1}{2}(u^{(n)}_{11,\alpha} + u^{(n)}_{22,\alpha}) \tag{3.5}$$

holds. This assumption, meaning that there is either no dominance or the beneficial allele is (partially) dominant, implies that every $\overline{w}_\alpha$ is concave (Bürger, 2009c). Hence, F is concave.

Remark 3.4. Corollary 3.3(c) does not apply if fitness contributions within loci are multiplicative. Then an unstable internal equilibrium may exist, and two monomorphic equilibria may be simultaneously stable (see the remark following the proof of Theorem 4.8 in Bürger, 2009c).

3.2. Deme-independent degree of intermediate dominance

Here, we make an assumption about the genetic architecture. Following N09b, we say there is deme-independent degree of intermediate dominance (DIDID) if

$$u^{(n)}_{i_n j_n,\alpha} = \vartheta^{(n)}_{i_n j_n} u^{(n)}_{i_n i_n,\alpha} + \vartheta^{(n)}_{j_n i_n} u^{(n)}_{j_n j_n,\alpha} \tag{3.6a}$$

holds for constants $\vartheta^{(n)}_{i_n j_n}$ such that

$$0 \leq \vartheta^{(n)}_{i_n j_n} \leq 1 \quad \text{and} \quad \vartheta^{(n)}_{j_n i_n} = 1 - \vartheta^{(n)}_{i_n j_n} \tag{3.6b}$$

for every α, every n, and every pair i_n, j_n. In particular, $\vartheta^{(n)}_{i_n i_n} = \frac{1}{2}$.

Obviously, DIDID covers complete dominance or recessiveness ($\vartheta^{(n)}_{i_n j_n} = 0$ or $= 1$ if $i_n \neq j_n$), and no dominance ($\vartheta^{(n)}_{i_n j_n} = \frac{1}{2}$), but not multiplicativity. We also note that DIDID includes the biologically important case of absence of genotype-by-environment interaction. In general, $F(\rho)$ is not concave under DIDID. Nevertheless, Theorem 3.14 in N09b establishes that under DIDID there exists exactly one stable gene-frequency equilibrium (point or manifold), and it is globally attracting. If an internal gene-frequency equilibrium exists, it is globally asymptotically stable.

Let us assume DIDID and that $\hat{\rho}$ is an internal gene-frequency equilibrium. It can be shown that

$$\hat{\xi}^{(n)}_{i_n j_n} = \hat{\bar{\xi}}^{(n)} \tag{3.7}$$

holds for every n and for every i_n and j_n (Appendix A). Therefore, (3.3) and (2.9) yield $W_{jk} = \sum_n \hat{\xi}^{(n)} = 1$. Hence, on $\hat{\Lambda}$, the dynamics of gamete frequencies reduces to that under a pure recombination process,

$$p_i' = \sum_{j,k} R_{i,jk} p_j p_k, \tag{3.8}$$

for which convergence to linkage equilibrium is well known (Geiringer, 1944). Of course, it is sufficient to consider internal gene-frequency equilibria because, otherwise, restriction to the subsimplex that supports $\hat{\rho}$ yields the result. Under pure recombination, the linkage-equilibrium manifold Ξ is globally attracting at a uniform geometric rate. For generic initial conditions the rate of approach is $1 - r_{\min}$, where $r_{\min}$ is the smallest two-locus recombination rate (see Lyubich, 1992; Nagylaki, 1993; Nagylaki et al., 1999). Therefore, the assumption $r_{\min} > 0$, which is weaker than (2.21), is sufficient to establish convergence to linkage equilibrium. We summarize these results as follows.

Theorem 3.5. *Suppose there is no epistasis and the assumptions (3.6), (2.19), and $r_{\min} > 0$ are satisfied. Then there exists a unique gene-frequency equilibrium $\hat{\rho}$ such that $\hat{p} = (\hat{\rho}, 0) \in \Lambda \cap \Xi$ is a globally asymptotically stable equilibrium point of (2.2).*

Remark 3.6. Because convergence of (3.8) to the linkage-equilibrium manifold Ξ occurs at a geometric rate $(1 - r_{\min})$, under the assumptions of Theorem 3.5 an equilibrium $\hat{p} = (\hat{\rho}, 0)$ of (2.2) is hyperbolic if $\hat{\rho}$ is a hyperbolic equilibrium of (2.16).

It would be interesting to identify the dominance patterns that lead to (3.7), hence, to (3.8). The set of these dominance patterns is a proper subset of all dominance patterns because for the internal equilibrium computed in Theorem 4.3 in Bürger (2009c), in general, the first locus does not satisfy (3.7).

4. The haploid Levene model

Here we consider a species that is haploid but reproduces sexually with recombination. General properties of the multilocus dynamics with viability selection in a panmictic population of haploids were derived by Kirzhner and Lyubich (1997). Various aspects of the one-locus haploid Levene model were studied by Strobeck (1979); see also Nagylaki and Lou (2008). Wiehe and Slatkin (1998) and Barton (2010) investigated haploid multilocus Levene models with certain forms of epistasis. As in the above-treated diploid case, and as in Wiehe and Slatkin (1998) and in one of Barton's (2010) models, we assume that the life cycle consists of viability selection, dispersal, and recombination. Moreover, soft selection is assumed.

All unexplained notation is as in Section 2. In haploids, the constant fitness $w_{i,\alpha}$ is assigned to gamete i in deme α. The mean fitness in deme α is simply $\overline{w}_\alpha = \sum_\alpha w_{i,\alpha} p_i$. Because recombination may occur between haplotypes originating from different demes, the recursion relations for the gamete frequencies are given by

$$p_i' = \sum_{j,k,\alpha,\beta} R_{i,jk} \frac{c_\alpha p_j w_{j,\alpha}}{\overline{w}_\alpha} \frac{c_\beta p_k w_{k,\beta}}{\overline{w}_\beta}. \tag{4.1}$$

From now on we assume absence of epistasis and set

$$w_{i,\alpha} = \sum_n u_{i_n,\alpha}^{(n)}, \tag{4.2a}$$

$$\bar{u}_\alpha^{(n)} = \sum_{i_n} u_{i_n,\alpha}^{(n)} p_{i_n}^{(n)}, \tag{4.2b}$$

$$\overline{w}_\alpha = \sum_i w_{i,\alpha} p_i = \sum_n \bar{u}_\alpha^{(n)}. \tag{4.2c}$$

Furthermore, we define

$$\xi_{i_n}^{(n)} = \sum_\alpha \frac{c_\alpha}{\overline{w}_\alpha} u_{i_n,\alpha}^{(n)} \tag{4.3}$$

and observe that, with $\bar{\xi}^{(n)}$ as in (2.8c), (2.9) continues to hold, i.e.,

$$\sum_n \bar{\xi}^{(n)} = 1. \tag{4.4}$$

Without epistasis we have

$$\sum_\alpha c_\alpha \frac{w_{i,\alpha}}{\overline{w}_\alpha} = \sum_\alpha \frac{c_\alpha}{\overline{w}_\alpha} \sum_n u_{i_n,\alpha}^{(n)} = \sum_n \xi_{i_n}^{(n)}, \tag{4.5}$$

and (4.1) simplifies to

$$p_i' = \sum_{j,k} R_{i,jk} p_j p_k \sum_n \xi_{j_n}^{(n)} \sum_l \xi_{k_l}^{(l)}. \tag{4.6}$$

Essentially the same, sometimes slightly simpler, calculations as in Sections 2 and 3 of N09b prove that the dynamics of gene frequencies has the form (2.10) and that Result 2.1 and Remarks 2.2 and 2.3 hold. An argument analogous to that yielding Corollary 3.8 in N09b shows that there exists exactly one stable gene-frequency equilibrium (point or manifold), and it is globally attracting. If an internal gene-frequency equilibrium exists, it is globally asymptotically stable.

For the rest of this section we assume (2.19). The results about the gene-frequency dynamics ensure that every $\rho(t)$ converges to some $\hat{\rho} \in \Omega$. Hence, for every solution $p(t)$ of (4.1), there exists a $\hat{\rho}$ such that the limit points of $p(t)$ are contained in $\hat{\Lambda}$. (For generic initial conditions solutions are attracted by the set $\hat{\Lambda}$ that is generated by the (unique) maximum $\hat{\rho}$ of geometric-mean fitness $\tilde{w}$.) For a given internal gene-frequency equilibrium $\hat{\rho}$, the equilibrium condition (2.15) together with (4.4) yields

$$\sum_n \hat{\xi}_{i_n}^{(n)} = \sum_n \hat{\xi}^{(n)} = 1. \tag{4.7}$$

Therefore (4.6) reduces to

$$p_i' = \sum_{j,k} R_{i,jk} p_j p_k, \tag{4.8}$$

which is the well-known dynamics of gamete frequencies under a pure recombination process. For (4.8) convergence to linkage equilibrium is well known and requires only $r_{\min} > 0$ instead of (2.21) (see Section 3.2). Thus, we have proved the following result:

Theorem 4.1. *Suppose there is no epistasis and the assumptions (2.19) and $r_{\min} > 0$ are satisfied. Then there exists a unique gene-frequency equilibrium $\hat{\rho}$ such that $\hat{p} = (\hat{\rho}, 0) \in \Lambda \cap \Xi$ is a globally asymptotically stable equilibrium point of (4.1).*

An analog of Remark 3.6 applies. We also note that the haploid model is not equivalent to the diploid model without epistasis and multiplicative fitnesses within loci. Instead, it is equivalent the diploid model with additive fitnesses within gametes and multiplicative fitnesses between them (see Remark 3.4 and Kirzhner and Lyubich, 1997).

5. Maintenance of polymorphism

In this section, we study the maintenance of multilocus polymorphism in the Levene model. In addition to assuming absence of epistasis, we assume that the degree of dominance is intermediate at every locus and in every deme. Thus, overdominance and underdominance are excluded. In a panmictic population, no polymorphism is possible in the absence of epistasis if there is intermediate dominance (Bürger, 2009b, Proposition 3.2 and Corollary 3.4). Although the Levene model lacks population structure, the following result shows that, nevertheless, it harbors the potential for extensive multilocus polymorphism under such conditions:

Result 5.1. *Assume an arbitrary number of multiallelic loci, $\Gamma \geq 2$ demes, and let all recombination rates be positive and fixed.*

Then there exists a nonempty open set of parameters such that for every parameter combination in this set, there is a unique, internal, asymptotically stable equilibrium point $\hat{p}$ of the dynamics (2.2). This equilibrium is in linkage equilibrium and it is globally asymptotically stable.

Essentially, this result was proved in Bürger (2009b, Theorem 2.2, Remark 2.3 (iii), Remark 2.4). There, only convergence to quasi-linkage equilibrium was shown. Theorem 3.1 demonstrates convergence to linkage equilibrium. More generally, Theorem 2.2 in Bürger (2009b) shows that Result 5.1 holds for arbitrary ergodic, i.e., irreducible and aperiodic, migration patterns, and not only for the Levene model. Such extensive polymorphism can occur if selection is weak relative to migration and recombination. The constructive proof in Bürger (2009a,b) requires balancing selection and a certain form of *average* overdominance across demes that can be achieved only if the direction of selection and the degree of dominance vary among demes.

If, however, the degree of dominance is not only intermediate within demes, but there is DIDID, generically, at most $\Gamma - 1$ diallelic loci can segregate (Proposition 3.18 in N09b). If, in addition, selection is sufficiently weak, then no polymorphism can be maintained (Proposition 2.6 and Remark 2.7 in Bürger, 2009b). Obviously, these results are in sharp contrast to Result 5.1 and underline that dominance plays an important role in maintaining genetic variation in a subdivided population.

For the rest of this section, we assume diallelic loci. Our main aim is to prove that the above mentioned upper bound $\Gamma - 1$ is indeed assumed on an open set of parameters. As a consequence, such polymorphic equilibria are structurally stable. We also show that if $L \geq \Gamma$ and if, as is nongeneric but may occur under additional constraints on the parameters, there exists an internal equilibrium, then it is a manifold.

For diallelic loci, we write $p^{(n)}$ for the frequency of $A_1^{(n)}$, and $1 - p^{(n)}$ for the frequency of $A_2^{(n)}$. The vector of gene frequencies ρ can then simply be represented as

$$\rho = (p^{(1)}, \ldots, p^{(L)})^T \in \Omega = [0, 1]^L, \tag{5.1}$$

where the definition of Ω is modified. The condition (2.15) for internal gene-frequency equilibria reduces to

$$\xi_1^{(n)} - \xi_2^{(n)} = \sum_\alpha \frac{c_\alpha}{\overline{w}_\alpha}(u_{1,\alpha}^{(n)} - u_{2,\alpha}^{(n)}) = 0 \quad \text{for every } n, \tag{5.2}$$

where the fitness contributions of the two alleles at locus n in deme α become

$$u_{1,\alpha}^{(n)} = p^{(n)}u_{11,\alpha}^{(n)} + (1 - p^{(n)})u_{12,\alpha}^{(n)} \tag{5.3a}$$

and

$$u_{2,\alpha}^{(n)} = p^{(n)}u_{12,\alpha}^{(n)} + (1 - p^{(n)})u_{22,\alpha}^{(n)}. \tag{5.3b}$$

The contribution to mean fitness of locus n in deme α is

$$\bar{u}_\alpha^{(n)} = (p^{(n)})^2 u_{11,\alpha}^{(n)} + 2p^{(n)}(1 - p^{(n)})u_{12,\alpha}^{(n)} + (1 - p^{(n)})^2 u_{22,\alpha}^{(n)}. \tag{5.4}$$

The allele-frequency dynamics, (2.16) or (2.17), simplifies to

$$\Delta p^{(n)} = p^{(n)}(1 - p^{(n)}) \sum_\alpha \frac{c_\alpha}{\overline{w}_\alpha}(u_{1,\alpha}^{(n)} - u_{2,\alpha}^{(n)}) \tag{5.5a}$$

$$= \frac{1}{2}p^{(n)}(1 - p^{(n)})\frac{\partial F(\rho)}{\partial p^{(n)}}. \tag{5.5b}$$

Remark 5.2. From (5.3), we obtain

$$u_{1,\alpha}^{(n)} - u_{2,\alpha}^{(n)} = p^{(n)}(u_{11,\alpha}^{(n)} - u_{12,\alpha}^{(n)}) + (1 - p^{(n)})(u_{12,\alpha}^{(n)} - u_{22,\alpha}^{(n)}). \tag{5.6}$$

Therefore, $u_{11,\alpha}^{(n)} \geq u_{12,\alpha}^{(n)} \geq u_{22,\alpha}^{(n)}$ and $u_{11,\alpha}^{(n)} > u_{22,\alpha}^{(n)}$ yield $u_{1,\alpha}^{(n)} - u_{2,\alpha}^{(n)} > 0$ if $0 < p^{(n)} < 1$. If $u_{1,\alpha}^{(n)} - u_{2,\alpha}^{(n)} > 0$ for every α, then (5.5a) and Remark 2.3 imply that allele $A_1^{(n)}$ becomes fixed. This confirms the intuition that, in the absence of overdominance and underdominance, an allele that is the best in every deme is fixed. Hence, a locus can be polymorphic only if each of the alleles is the best in at least one deme.

For two alleles, also the definition of DIDID simplifies. There is DIDID if for every $n \in \mathsf{L}$, there exist constants $\vartheta^{(n)}$ such that for every $\alpha \in \mathsf{G}$,

$$u_{12,\alpha}^{(n)} = \vartheta^{(n)}u_{11,\alpha}^{(n)} + (1 - \vartheta^{(n)})u_{22,\alpha}^{(n)} \tag{5.7a}$$

holds, where

$$0 \leq \vartheta^{(n)} \leq 1. \tag{5.7b}$$

If (5.7) holds, the condition (5.2) for an internal gene-frequency equilibrium simplifies to

$$\xi_{11}^{(n)} - \xi_{22}^{(n)} = \sum_\alpha c_\alpha \frac{u_{11,\alpha}^{(n)} - u_{22,\alpha}^{(n)}}{\overline{w}_\alpha} = 0 \quad \text{for every } n \in \mathsf{L}, \tag{5.8}$$

because

$$u_{1,\alpha}^{(n)} - u_{2,\alpha}^{(n)} = [\vartheta^{(n)} + p^{(n)}(1 - 2\vartheta^{(n)})](u_{11,\alpha}^{(n)} - u_{22,\alpha}^{(n)}), \tag{5.9}$$

and $\vartheta^{(n)} + p^{(n)}(1 - 2\vartheta^{(n)}) > 0$ if $0 < p^{(n)} < 1$. We shall write $\vartheta = (\vartheta^{(1)}, \ldots, \vartheta^{(L)})^T \in [0, 1]^L$. Therefore, the dynamics of gamete frequencies (2.2) is uniquely determined by the linkage map and by the parameters $c_1, \ldots, c_{\Gamma-1}$, ϑ, $u_{11,\alpha}^{(n)}$ and $u_{22,\alpha}^{(n)}$, where $\alpha \in \mathsf{G}$ and $n \in \mathsf{L}$.

Our main result in this section is the following.

Theorem 5.3. *Assume $L \leq \Gamma - 1$ diallelic loci with DIDID, i.e., (5.7). Let $\vartheta \in [0, 1]^L$ be arbitrary but fixed. Then there exists an open nonempty set U of parameters such that for every parameter combination in U, there is a unique, internal, asymptotically stable equilibrium point $\hat{p}$ of (2.2). This equilibrium is in linkage equilibrium and globally attracting. The set U is independent of the choice of the recombination rates.*

Together with Proposition 3.18 in N09b, this result implies that under DIDID and if $L \geq \Gamma$, at most $\Gamma - 1$ diallelic loci can be maintained polymorphic for an open set of parameters. The proof of the above theorem is based on

Proposition 5.4. *Under the assumptions of Theorem 5.3 there exists an open, nonempty subset U of the parameter space such that for every element in U there is an isolated, internal, asymptotically stable equilibrium point $\hat{\rho}$ of the gene-frequency dynamics (5.5a). It is the unique solution of (5.8) in $\operatorname{int}\Omega = (0, 1)^L$.*

The basic idea for the proof of Proposition 5.4 is to find parameter combinations such that $(\frac{1}{2}, \ldots, \frac{1}{2})^T \in \mathbb{R}^L$ is an equilibrium of the gene-frequency dynamics (5.5a), and then to show that every sufficiently small perturbation of the parameters still yields an (isolated) internal equilibrium. The proof of this proposition is rather technical and constructs the set U, which is the same in Theorem 5.3 and Proposition 5.4. A more detailed formulation of the proposition and its proof are given in the Appendix.

Proof of Theorem 5.3. Statements (a) and (b) of Corollary 3.3 show that for every $\hat{\rho}$ in the open set U of Proposition 5.4, $\hat{p} = (\hat{\rho}, 0)$ is the desired unique and internal equilibrium point of (2.2). Because the proof of Proposition 5.4 is based entirely on the gene-frequency dynamics (5.5a), the construction of U is independent of the linkage map. □

Now we formulate and prove our second main result of this section.

Theorem 5.5. *Assume $L \geq \Gamma$ diallelic loci with DIDID. If there exists an internal equilibrium $\hat{\rho}$ of the gene-frequency dynamics (5.5a), then there is a manifold of equilibrium points containing it. Generically, this manifold has dimension $L - \Gamma + 1$.*

Proof. Let $V = (V_{n\alpha})$ denote the $L \times \Gamma$ matrix with entries $V_{n\alpha} = u^{(n)}_{11,\alpha} - u^{(n)}_{22,\alpha}$. By assumption, there exists an internal equilibrium $\hat{\rho}$ of the gene-frequency dynamics (5.5a). If we write $\hat{\bar{w}}_\alpha = \bar{w}_\alpha(\hat{\rho})$, (5.8) informs us that

$$\sum_\alpha V_{n\alpha}\frac{c_\alpha}{\hat{\bar{w}}_\alpha} = \sum_\alpha \frac{c_\alpha}{\hat{\bar{w}}_\alpha}(u^{(n)}_{11,\alpha} - u^{(n)}_{22,\alpha}) = 0, \quad n \in \mathsf{L}, \tag{5.10}$$

holds. In matrix form this reads

$$V\hat{x} = 0, \tag{5.11}$$

where $\hat{x} = (c_1/\hat{\bar{w}}_1, \ldots, c_\Gamma/\hat{\bar{w}}_\Gamma)^T$. Generically, we therefore have $\dim(\ker V) = 1$ and, because $L \geq \Gamma$,

$$\operatorname{rank} V^T = \operatorname{rank} V = \min\{L, \Gamma\} - 1 = \Gamma - 1. \tag{5.12}$$

Although scalar multiples of $\hat{x}$ also solve (5.11), scalar multiples of $(\hat{\bar{w}}_1, \ldots, \hat{\bar{w}}_\Gamma)^T$ do not give rise to further solutions of (5.10) because, by Result 2.1, geometric-mean fitness $\tilde{w}$ is maximized, hence constant, at gene-frequency equilibria.

Therefore, it is sufficient to determine the dimension of the solution space of the system

$$\bar{w}_\alpha(\rho) = \hat{\bar{w}}_\alpha, \quad \alpha \in \mathsf{G}, \tag{5.13}$$

for the given vector $(\hat{\bar{w}}_\alpha)$. This is a system of Γ quadratic equations in the L unknowns $p^{(n)}$. We transform it into a system of Γ linear equations in the L unknowns $p^{(n)}$ by using the nonlinear transformation $f = (f^{(1)}, \ldots, f^{(L)})^T$ defined by

$$z^{(n)} = f^{(n)}(p^{(n)}) = 2\vartheta^{(n)}p^{(n)} + (1 - 2\vartheta^{(n)})(p^{(n)})^2. \tag{5.14}$$

This is the specification of (3.20) in N09b for diallelic loci. By Lemma 3.1 in Nagylaki (2009a), every $f^{(n)}$ is a homeomorphism on $[0, 1]$. Denoting $z = (z^{(1)}, \ldots, z^{(L)})^T$, we obtain from (3.23) in N09b by a simple calculation,

$$\begin{aligned}\bar{w}_\alpha(\rho) &= \bar{W}_\alpha(z) = \sum_n u^{(n)}_{22,\alpha} + \sum_n (u^{(n)}_{11,\alpha} - u^{(n)}_{22,\alpha})z^{(n)} \\ &= \sum_n u^{(n)}_{22,\alpha} + \sum_n V_{n\alpha}z^{(n)},\end{aligned} \tag{5.15}$$

which is linear in z for every $\alpha \in \mathsf{G}$ (cf. Lemma 3.13 in N09b). Therefore, the system (5.13) of quadratic equations in ρ is equivalent to the linear system

$$\bar{W}_\alpha(z) = \hat{\bar{w}}_\alpha, \quad \alpha \in \mathsf{G}, \tag{5.16}$$

in z of the same dimension. This can be written as

$$V^T z = b, \tag{5.17}$$

where $b = (b_1, \ldots, b_\Gamma)^T$ and $b_\alpha = \hat{\bar{w}}_\alpha - \sum_n u^{(n)}_{22,\alpha}$. Because (5.17) has the nontrivial solution $\hat{z} = f(\hat{\rho})$, standard linear algebra shows that the dimension of its solution space, hence that of (5.13), generically equals

$$\dim(\ker V^T) = L - \operatorname{rank} V^T = L - (\Gamma - 1) = L - \Gamma + 1. \tag{5.18}$$

Here, the second equality follows from (5.12) which holds generically. □

Because under the conditions of the above theorem a manifold of gene-frequency equilibria exists, it cannot be inferred that convergence to linkage equilibrium occurs although this seems likely. The reason is that the application of Theorem 3.1 requires isolated gene-frequency equilibria or at least a proof that every $\rho(t)$ converges (to a single point); cf. Remark 3.2.

6. A quantitative-genetic model

We apply the above results and those of N09b to a simple quantitative-genetic model and discuss their implications. To this end, we consider a quantitative trait that is determined additively (i.e., without epistasis) by $L \geq 2$ multiallelic loci. Intermediate dominance and genotype-by-environment ($G \times E$) interaction are admitted. Let the multilocus genotype ij have the trait value

$$X = X_{ij,\alpha} = \sum_n \gamma^{(n)}_{i_n j_n,\alpha} \tag{6.1}$$

in deme α, where $\gamma^{(n)}_{i_n j_n,\alpha}$ is the effect on the trait in deme α assigned to the single-locus genotype $A^{(n)}_{i_n}A^{(n)}_{j_n}$. To avoid degeneracy, we posit $\gamma^{(n)}_{i_n i_n,\alpha} \neq \gamma^{(n)}_{j_n j_n,\alpha}$ if $i_n \neq j_n$.

We assume that the trait is under linear directional selection in each of the $\Gamma \geq 1$ demes, i.e.,

$$W_\alpha(X) = 1 - s_\alpha X, \tag{6.2}$$

where the range of possible values X and the selection coefficients $s_\alpha \neq 0$ are constrained such that $W_\alpha(X) > 0$ for every $\alpha \in \mathsf{G}$. Therefore, we define the single-locus contribution of $A^{(n)}_{i_n}A^{(n)}_{j_n}$ to fitness by

$$u^{(n)}_{i_n j_n,\alpha} = \frac{1}{L} - s_\alpha \gamma^{(n)}_{i_n j_n,\alpha}. \tag{6.3}$$

Then, as in (2.4),

$$w_{ij,\alpha} = W_\alpha(X_{ij,\alpha}) = \sum_n u^{(n)}_{i_n j_n,\alpha}. \tag{6.4}$$

A glance at (6.3) reveals that, for given L, Γ, and selection coefficients s_α, (6.3) establishes a one-to-one correspondence between the single-locus fitness contributions $u^{(n)}_{i_n j_n,\alpha}$ and the single-locus trait effects $\gamma^{(n)}_{i_n j_n,\alpha}$.

Therefore, this model of linear selection on a quantitative trait is as general as the nonepistatic model introduced in Section 2 and studied in Sections 3 and 5. Obviously, no generality would be lost by setting $s_\alpha = 1$ for every α. However, we refrain from doing so. Moreover, it is immediate from (6.3) that no or intermediate dominance at the trait level is equivalent to no or intermediate dominance, respectively, at the fitness level. The dominance relations are reversed in demes with $s_\alpha > 0$.

In analogy to (3.6), we say there is DIDID on the trait level if there exist constants $\vartheta^{(n)}_{i_n j_n}$ such that

$$\gamma^{(n)}_{i_n j_n,\alpha} = \vartheta^{(n)}_{i_n j_n}\gamma^{(n)}_{i_n i_n,\alpha} + \vartheta^{(n)}_{j_n i_n}\gamma^{(n)}_{j_n j_n,\alpha} \tag{6.5a}$$

and

$$0 \leq \vartheta^{(n)}_{i_n j_n} \leq 1 \quad \text{and} \quad \vartheta^{(n)}_{j_n i_n} = 1 - \vartheta^{(n)}_{i_n j_n} \tag{6.5b}$$

for every α, every n, and every pair i_n, j_n. Trivially, the relation (6.3) transforms DIDID on the trait level to DIDID on the fitness level, and vice versa; the constants $\vartheta^{(n)}_{i_n j_n}$ are the same for trait and fitness.

Genotype-by-environment interaction is absent on the trait level if there exist constants $\gamma^{(n)}_{i_n j_n}$ and γ_α such that

$$\gamma^{(n)}_{i_n j_n,\alpha} = \gamma^{(n)}_{i_n j_n} + \gamma_\alpha \tag{6.6}$$

holds for every α, every n, and every pair i_n, j_n. Because we assume intermediate dominance at the trait level, (6.6) implies the existence of constants $\vartheta^{(n)}_{i_n j_n}$ satisfying (6.5b) such that

$$\gamma^{(n)}_{i_n j_n} = \vartheta^{(n)}_{i_n j_n}\gamma^{(n)}_{i_n i_n} + \vartheta^{(n)}_{j_n i_n}\gamma^{(n)}_{j_n j_n}. \tag{6.7}$$

Therefore, (6.6) implies (6.5). Summarizing, there is no $G \times E$ interaction on the trait level if and only if there is DIDID and (6.6) holds for all homozygous single-locus effects. A further simple consideration establishes equivalence of absence of $G \times E$ interaction at the trait and the fitness level.

An important consequence of this one-to-one relation between fitness and trait value is that results that hold generically for the model defined in terms of the fitness contributions $u^{(n)}_{i_n j_n,\alpha}$ also

hold generically for this model of a quantitative trait under linear selection. Especially, Result 5.1 and Theorems 5.3 and 5.5 apply. Thus, if the number of demes is at least two, for an open set of parameters a globally asymptotically stable equilibrium exists such that

(i) an arbitrary number of multiallelic loci is polymorphic;
(ii) up to $\Gamma - 1$ diallelic loci are polymorphic if there is DIDID.

Parameter combinations yielding (i) can be constructed by applying the procedure in Proposition 2.1 in Bürger (2009b) with $m_2 = 1 - m_1 = 1 - c_2 = c_1$ and substituting $-s_\alpha \gamma^{(n)}_{i_n j_n,\alpha}$ for the parameters $u^{(n)}_{i_n j_n,\alpha}$ used there (here, m_α is the backward migration rate into α). In fact, it is easy to show that this is possible if $\gamma^{(n)}_{i_n i_n,\alpha} = \gamma^{(n)}_{i_n i_n}$, i.e., in the absence of interactions between homozygous effects and environment.

An unresolved question concerns the number of alleles that can be maintained at a locus if there is DIDID. In N09b (Corollary 3.9) it is shown that if $\bar{I} = \frac{1}{L}\sum_n I_n > \Gamma$, then generically no internal gene-frequency equilibrium exists. Therefore, if the number of alleles is the same at every locus ($I_n = I$ for every n), the number of demes is a generic upper bound on the number I of alleles per locus. Clearly, a generalization of (ii) to multiallelic loci would be desirable.

Example 6.1. We specialize the model (6.1)–(6.4) by assuming absence of $G \times E$ interaction, i.e., (6.6). Therefore, (6.5) holds and we define

$$\vartheta^{(n)}_{i_n} = \sum_{j_n} \vartheta^{(n)}_{i_n j_n} p^{(n)}_{j_n}. \tag{6.8}$$

From (2.7) and (2.6) we obtain by straightforward calculations employing (6.3), (6.5) and (6.6):

$$u^{(n)}_{i_n,\alpha} = \frac{1}{L} - s_\alpha \gamma_\alpha - s_\alpha \left(\gamma^{(n)}_{i_n i_n} \vartheta^{(n)}_{i_n} + \sum_{j_n} \vartheta^{(n)}_{j_n i_n} \gamma^{(n)}_{j_n j_n} p^{(n)}_{j_n} \right), \tag{6.9a}$$

$$\bar{u}^{(n)}_{\alpha} = \frac{1}{L} - s_\alpha \gamma_\alpha - 2 s_\alpha \sum_{j_n} \vartheta^{(n)}_{j_n} \gamma^{(n)}_{j_n j_n} p^{(n)}_{j_n}, \tag{6.9b}$$

and

$$u^{(n)}_{i_n,\alpha} - \bar{u}^{(n)}_{\alpha} = s_\alpha \phi^{(n)}_{i_n}, \tag{6.10}$$

where

$$\phi^{(n)}_{i_n} = \sum_{j_n} (2\vartheta^{(n)}_{j_n} - \vartheta^{(n)}_{j_n i_n}) \gamma^{(n)}_{j_n j_n} p^{(n)}_{j_n} - \vartheta^{(n)}_{i_n} \gamma^{(n)}_{i_n i_n}. \tag{6.11}$$

Therefore, the condition (2.15) for internal gene-frequency equilibria becomes

$$\phi^{(n)}_{i_n} \sum_\alpha \frac{c_\alpha s_\alpha}{\overline{w}_\alpha} = 0 \tag{6.12}$$

for every n and i_n. We conjecture that, at least generically, for every n, $\phi^{(n)}_{i_n} \neq 0$ holds for at least one i_n (in fact, for all but one). If this is true, then

$$\sum_\alpha \frac{c_\alpha s_\alpha}{\overline{w}_\alpha} = 0 \tag{6.13}$$

is a necessary and sufficient condition for the existence of an internal gene-frequency equilibrium in the absence of $G \times E$ interaction.

In the absence of dominance, this is easy to prove because we have $\vartheta^{(n)}_{i_n} = \vartheta^{(n)}_{i_n j_n} = \frac{1}{2}$ for every n and every i_n, j_n, hence

$$\phi^{(n)}_{i_n} = \frac{1}{2} \left(\sum_{j_n} \gamma^{(n)}_{j_n j_n} p^{(n)}_{j_n} - \gamma^{(n)}_{i_n i_n} \right). \tag{6.14}$$

Obviously, for each n, $\phi^{(n)}_{i_n} = 0$ can hold for at most one i_n. Thus, (6.13) is the condition for an internal gene-frequency equilibrium.

Example 6.2. Here, we specialize to diallelic loci and use the notation introduced in Sections 2 and 5, e.g., $p^{(n)} = p^{(n)}_1$ and $\vartheta = \vartheta_{12}$. By Corollary 3.3, the equilibrium and stability structure can be determined by studying the relatively simple system (5.5a), where a straightforward calculation yields

$$u^{(n)}_{1,\alpha} - u^{(n)}_{2,\alpha} = s_\alpha [\gamma^{(n)}_{22,\alpha} - \gamma^{(n)}_{12,\alpha} + p^{(n)} (2\gamma^{(n)}_{12,\alpha} - \gamma^{(n)}_{11,\alpha} - \gamma^{(n)}_{22,\alpha})]. \tag{6.15}$$

In the absence of $G \times E$ interaction, i.e., if (6.6) holds, (6.15) simplifies to

$$u^{(n)}_{1,\alpha} - u^{(n)}_{2,\alpha} = s_\alpha (\gamma^{(n)}_{22} - \gamma^{(n)}_{11}) [\vartheta + p^{(n)} (1 - 2\vartheta)]. \tag{6.16}$$

Because $\vartheta + p^{(n)}(1 - 2\vartheta) > 0$ if $0 < p^{(n)} < 1$, the condition (5.2) for an internal gene-frequency equilibrium again collapses to the single equation (6.13) in the L unknowns $p^{(n)}$.

Because $\overline{w}_\alpha$ is quadratic in $p^{(n)}$, (6.13) is equivalent to a single polynomial equation (of degree $2\Gamma - 2$) in the L variables $p^{(n)}$. Therefore, if an internal solution $\hat{\rho}$ exists, there is a manifold of solutions of (generic) dimension $L - 1$. For more than two demes, this exceeds the dimension $L - \Gamma + 1$ derived in Theorem 5.5. As a consistency check, we observe that the matrix V defined in the proof of Theorem 5.5 has the entries $V_{n\alpha} = s_\alpha \phi^{(n)}_1 / (1 - p^{(n)})$ because $\phi^{(n)}_1 = (1 - p^{(n)})(\gamma^{(n)}_{22} - \gamma^{(n)}_{11})[\vartheta + p^{(n)}(1 - 2\vartheta)]$. Therefore, it has rank 1 and a calculation as in (5.18) yields $L - 1$. Thus, if $\Gamma > 2$, the equilibrium structure in the absence of $G \times E$ interaction is even more highly degenerate than under DIDID alone.

The above examples demonstrate that (6.13) is a necessary and sufficient condition for an internal gene-frequency equilibrium if there is no $G \times E$ interaction and either there is no dominance or loci are diallelic. In the absence of $G \times E$ interaction, an internal equilibrium cannot exist if all selection coefficients s_α have the same sign (for a more general statement, see Remark 5.2). If, for instance, $s_\alpha > 0$ for every α, then the gamete with the largest genotypic value has maximum fitness and becomes fixed (cf. N09b, Proposition 3.15). If the selection coefficients s_α vary in sign, it is always possible to choose the relative demes sizes c_α such that (6.13) has an internal solution $\hat{\rho}$, hence a manifold of solutions. In fact, there is an open set of such parameters combinations $(c_1, \ldots, c_{\Gamma-1})$. For other deme sizes, there is at least one isolated, asymptotically stable equilibrium. At such a stationary state, at most $\Gamma - 1$ loci can be polymorphic.

We summarize the main results of this section.

Corollary 6.3. *Assume diallelic loci and let all recombination rates be positive and fixed. Further, assume the selection model given by (6.1)–(6.4), i.e., linear directional selection on a quantitative trait that is determined additively by L diallelic loci exhibiting intermediate dominance.*

(a) *There is an open subset of the full parameter space for which a globally asymptotically stable internal, hence fully polymorphic, equilibrium exists.*
(b) *If there is DIDID, then, generically, at most $\Gamma - 1$ loci can be polymorphic at an equilibrium. An internal equilibrium exists for an open subset of parameters if and only if $L \leq \Gamma - 1$. If $L \geq \Gamma$ and an internal equilibrium exists, then it is a manifold of equilibria which, generically, has dimension $L - \Gamma + 1$.*
(c) *If there is no $G \times E$ interaction, then there is an open subset of parameters for which a unique internal equilibrium exists. If an internal equilibrium exists, it is the manifold of equilibria that solves Eq. (6.13). Generically, its dimension is $L - 1$.*
(d) *Every equilibrium is in linkage equilibrium.*

Among others, this result demonstrates that biologically reasonable but special assumptions can lead to nongeneric, and nonrobust, model behavior. Therefore, caution is necessary when generalizing conclusions obtained from specific models.

A detailed and rather complete analysis of the diallelic two-locus case in two demes was performed in Bürger (2009c).

7. Weak epistasis

Several of our results can be extended to weak epistasis. By weak epistasis we mean that there is a small $\epsilon > 0$, such that the fitness scheme has the form

$$w_{ij,\alpha} = \sum_n u^{(n)}_{i_n j_n,\alpha} + s_{ij,\alpha}, \tag{7.1a}$$

where

$$\left|s_{ij,\alpha}\right| < \epsilon \quad \text{for every } i, j, \text{ and } \alpha. \tag{7.1b}$$

First, we generalize Theorem 3.1 and show that global convergence to quasi-linkage equilibrium occurs under weak epistasis. Then we briefly point out some applications for the maintenance of polymorphism. Throughout, we assume (2.19)–(2.21).

7.1. Convergence

As noted by Nagylaki et al. (1999), the assumption of hyperbolicity of every equilibrium is not robust to small perturbations because limit sets need not change continuously. What has good behavior under perturbations is the set of chain-recurrent points, which contains the limit sets of all orbits (Conley, 1978; Akin, 1993). For the Levene model without epistasis, we have the following:

Lemma 7.1. *In the case of no epistasis, the only chain-recurrent points of (2.2) are its equilibria.*

Proof. By Result 2.1, the Lyapunov function F takes only finitely many values on the set Λ of gene-frequency equilibria. Therefore, Theorem 3.16 in Akin (1993) shows that every chain-recurrent point is contained in Λ. By our general assumption (2.19), Λ has finitely many components $\hat{\Lambda}$. Since, by (3.1), on $\hat{\Lambda}$ the dynamics reduces to the panmictic multilocus dynamics (3.2) with no epistasis, Lemma 2.2 of Nagylaki et al. (1999) yields the assertion. □

To formulate the main result of this section, we introduce the following notation. We write $\Sigma_0 \subset \Delta_J$ for the set of equilibria of (2.2) if there is no epistasis ($\epsilon = 0$), and we write $\Sigma_{w(\epsilon)} \subset \Delta_J$ for the set of equilibria for a fitness scheme $w_{ij,\alpha}$ that satisfies (7.1) with $\epsilon > 0$.

Theorem 7.2. *Let fitness contributions $u^{(n)}_{i_n j_n,\alpha}$ be given, such that without epistasis every equilibrium of (2.2) is hyperbolic. Further, let all recombination probabilities be fixed (and positive), and let $\epsilon > 0$ be sufficiently small. Then for every set of fitnesses $w_{ij,\alpha}$ satisfying (7.1) the following holds:*

(a) *The set of equilibria Σ_0 contains only isolated points, as does the set of equilibria $\Sigma_{w(\epsilon)}$. As $\epsilon \to 0$ in (7.1), each equilibrium in $\Sigma_{w(\epsilon)}$ converges to the corresponding equilibrium in Σ_0.*

(b) *In the neighborhood of each equilibrium point in Σ_0, there exists exactly one equilibrium point in $\Sigma_{w(\epsilon)}$. The stability of each equilibrium in $\Sigma_{w(\epsilon)}$ is the same as that of the corresponding equilibrium in Σ_0; i.e., each pair is either asymptotically stable or unstable.*

(c) *Every solution $p(t)$ of (2.2) converges to one of the equilibrium points in $\Sigma_{w(\epsilon)}$.*

Proof. Parts (a) and (b) are consequences of the implicit function theorem and the Hartman–Grobman theorem, and follow immediately from Theorem 4.4 in Karlin and McGregor (1972). The only issue that remains to be shown in (b) is that perturbed equilibria do not leave Δ_J (cf. Remark 2.3 in Nagylaki et al., 1999). This follows from the explicit characterization

$$p^{(n)}_{i_n} = 0 \quad \text{or} \quad \xi^{(n)}_{i_n} = \bar{\xi}^{(n)} \tag{7.2}$$

of equilibria if $\epsilon = 0$; see (2.16). Let $\hat{p}$ be some equilibrium for $\epsilon = 0$, and denote the set of alleles present at locus n by $J_n = \{i_n : \hat{p}^{(n)}_{i_n} > 0\}$. The face of Δ_J determined by $p_i = 0$ if $i \notin \prod_n J_n$ is invariant under the dynamics (2.2) for every $\epsilon > 0$, as can be seen from the alternative representation

$$p_i' = \sum_\alpha c_\alpha \left(p_i \frac{w_{i,\alpha}}{\overline{w}_\alpha} - D_{i,\alpha} \right), \tag{7.3}$$

where $D_{i,\alpha}$ is a measure of linkage disequilibrium in gamete i in deme α (see Eq. (4.47) in N09b). Therefore, the equilibrium $\hat{p}$, which is hyperbolic for $\epsilon = 0$, persists in this face for $\epsilon > 0$.

(c) Upon replacing the reference to Lemma 2.2 in the proof of Theorem 2.3 in Nagylaki et al. (1999) by one to the above Lemma 7.1, their proof applies unaltered. However, our assumption (7.1) is slightly weaker than their assumption (2.1). Our assumption is uniform in the sense that we first choose $\epsilon > 0$ and then admit all parameter combinations $w_{ij,\alpha}$ satisfying (7.1), whereas following their formulation, we first had to fix a set of $s_{ij,\alpha}$ and then $\epsilon > 0$. Because Corollary 32 in Akin (1993), which is used in the proof of Theorem 2.3 in Nagylaki et al. (1999), admits this greater generality, their proof indeed applies. □

Remark 7.3.

1. Hyperbolicity of equilibria of (2.2) with $\epsilon = 0$ is a generic property. For a single panmictic deme, this was proved by Nagylaki et al. (1999, Appendix A). In general, (2.2) shows that the Jacobian (at any point) is simply a linear combination of single-deme Jacobians.
2. With weak epistasis, equilibria may be in linkage disequilibrium, but deviate from Ξ only to order $O(\epsilon)$.
3. If we define $\tilde{w}(p) = \prod_\alpha [\overline{w}_\alpha(p)]^{c_\alpha}$ if $\epsilon > 0$, then $\Delta\tilde{w}(p) > 0$ holds if p is bounded away from the set of gene-frequency equilibria Λ.

We conjecture that an equilibrium $\hat{p} = (\hat{\rho}, 0)$ of (2.2) with $\epsilon = 0$ is hyperbolic if $\hat{\rho}$ is a hyperbolic equilibrium of (2.16). In the presence of selection, in general, this cannot be inferred from the results or proofs in Lyubich (1992) because it has not been shown that convergence to linkage equilibrium occurs at a geometric rate. If there are only two loci, or if there is DIDID, or if selection acts on haploids, then Theorem 4.6 in N09b, or Remark 3.6, or the remark following Theorem 4.1, respectively, establish the conjecture. It can also be established for three diallelic loci (Bürger, unpublished).

7.2. Maintenance of polymorphism

With the help of Theorem 7.2, several of the results on the maintenance of polymorphism can be generalized to weak epistasis. Of course, Result 5.1 generalizes to weak epistasis. The only difference then is that equilibria are not necessarily in linkage equilibrium but in quasi-linkage equilibrium. In fact, this was already proved in Bürger (2009b, Theorem 2.2 and Remark 2.4). Proposition 3.18 in N09b and Theorem 5.3 have the following generalization.

Theorem 7.4. *Assume weak epistasis and diallelic loci with DIDID.*

(a) *If $L \geq \Gamma$, at most $\Gamma - 1$ loci can be maintained polymorphic for an open set of parameters.*

(b) *If $L \leq \Gamma - 1$ and $\vartheta \in [0, 1]^L$ in (5.7) is arbitrary but fixed, then there exists an open nonempty set W of fitness schemes satisfying (7.1) and of deme proportions $(c_1, \ldots, c_{\Gamma-1})$, such that for every parameter combination in W, there is a*

unique, internal, asymptotically stable equilibrium point $\hat{p}$ *of* (2.2). *This equilibrium is in quasi-linkage equilibrium and globally attracting. The set* W *is independent of the choice of the recombination rates.*

Proof. (a) Proposition 3.18 in N09b shows that generically, $\operatorname{int}\Lambda = \emptyset$. Because, in the proof of Theorem 7.2(b), it was shown that boundary equilibria persist in the face where they exist for $\epsilon = 0$, Theorem 7.2 yields the assertion.
(b) This is a simple consequence of Theorems 5.3 and 7.2. □

Also statements (a) and (b) in Corollary 6.3 can be generalized to weak epistasis. In (d), quasi-linkage equilibrium can be stated.

8. Discussion

The analysis of multilocus systems is greatly simplified, and often only feasible, if linkage equilibrium or, at least, quasi-linkage equilibrium can be assumed (e.g. Karlin and Liberman, 1979; Turelli and Barton, 1990; Christiansen, 1999; Bürger, 2000). For the classical multilocus selection model, it has long been known that global convergence to linkage equilibrium occurs if there is no (additive) epistasis (Kun and Lyubich, 1979, 1980; Lyubich, 1992). If either epistasis or selection is weak, generic convergence to a stationary point in quasi-linkage equilibrium has been proved (Nagylaki, 1993; Nagylaki et al., 1999). The latter results can be generalized to subdivided populations as follows. If migration and epistasis are weak (relative to recombination) or if selection is weak (relative to migration and recombination), then global convergence to a stationary point in quasi-linkage equilibrium occurs generically (Bürger, 2009a). However, if migration is moderately strong, stable linkage disequilibrium may persist in the absence of epistasis (Li and Nei, 1974).

For the multilocus Levene model without epistasis, and under the generic assumption that all equilibria are isolated, Nagylaki (2009b) proved global convergence to an equilibrium point in linkage equilibrium if there are two (multiallelic) loci or if there is no dominance. He conjectured global convergence for multiple multiallelic loci, even if equilibrium points are not isolated. In Section 3, we establish this conjecture under the generic assumption of isolated gene-frequency equilibria (Theorem 3.1). For the haploid Levene model, a slightly stronger result is demonstrated in Section 4.

Whenever convergence to an equilibrium point in linkage equilibrium occurs in the Levene model without epistasis, the analysis of the model simplifies greatly. Then it is sufficient to study the system of recursion equations (2.16), which describes the dynamics of gene frequencies under the assumption of linkage equilibrium. Corollary 3.3 summarizes the most important conclusions that can be drawn from investigating (2.16) and its Lyapunov function $F = \ln \tilde{w}$ (2.14). Further useful properties of the gene-frequency dynamics may be found in Section 3 of N09b.

With the help of perturbation methods (Karlin and McGregor, 1972; Nagylaki et al., 1999), the convergence result Theorem 3.1 can be extended to weak epistasis. Then equilibria are not necessarily in linkage equilibrium, but they are in quasi-linkage equilibrium (Theorem 7.2). Therefore, many results that hold without epistasis can be extended to weak epistasis. For an example, see Theorem 7.4.

In Section 5, we apply our convergence results to investigate the maintenance of multilocus polymorphism. Overdominance and underdominance (within demes) are excluded. It has been shown previously that, if there is intermediate dominance in every deme, arbitrarily many multiallelic loci can be fully polymorphic on an open set of parameters (Result 5.1). In the proof, this open set was constructed by assuming, among others, that at each locus and in each deme, the fitter alleles are partially dominant. Because an internal equilibrium requires that the direction of selection is different in at least two demes, the proof of this result invokes $G \times E$ interaction. In N09b it was shown that for DIDID, generically, the number of segregating loci is strictly less than the number Γ of demes. Hence, some form of $G \times E$ interaction is necessary to maintain Γ or more loci polymorphic. Theorem 5.3 complements Nagylaki's result and shows that with DIDID, $\Gamma - 1$ loci can indeed be maintained at a stable equilibrium for an open set of parameters. If $L \geq \Gamma$ and, as is nongeneric, an internal equilibrium exists, Theorem 5.5 establishes that a manifold of equilibria exists. Generically, its dimension is $L - \Gamma + 1$. If there is DIDID and selection is sufficiently weak, then no polymorphism is possible in the Levene model without epistasis.

Of course, DIDID itself is a nongeneric property. Perturbations of isolated equilibrium points can be studied using, for instance, Karlin and McGregor's (1972) method of small parameters which requires that equilibria are hyperbolic. Thus, if there is an asymptotically stable equilibrium under DIDID, there will be an asymptotically stable equilibrium in its neighborhood if the deviation from DIDID is small. A much stronger result can be inferred from Lemma 7.1 and the proof of Theorem 7.2. If all equilibria are hyperbolic when there is DIDID, under small deviations from DIDID, the global dynamics will remain qualitatively unchanged.

In Section 6, these results are applied to a quantitative trait that is determined by L additive loci (i.e., without epistasis but with intermediate dominance). If the trait is under linear directional selection in every deme, the resulting model is equivalent to the general nonepistatic model studied in this paper. Hence, the results of Section 5 concerning the maintenance of variation apply unaltered. Interestingly, under the assumption of no $G \times E$ interaction on the trait level, there is an open set of parameters such that an internal equilibrium exists. However, because absence of $G \times E$ interaction implies DIDID, this is a manifold of equilibria. In fact, in this case, its dimension is $L - 1$ instead of $L - \Gamma + 1$. Thus, if there are more than two demes, the assumption of no $G \times E$ interaction in the Levene model with linear selection, leads to a highly degenerate equilibrium structure. In such a case, an arbitrarily small perturbation of the parameters that introduces $G \times E$ interaction, may drastically change the equilibrium structure and the dynamics (see Bürger, 2009c). As already mentioned in the Introduction, these results (Corollary 6.3) should serve as a warning when studying models under highly specialized assumptions, even if they are biologically well motivated.

The fact that the maximum possible number of polymorphic loci may be constrained by the number of demes demonstrates that, even in the absence of epistasis and in the presence of linkage equilibrium, one cannot simply extrapolate single-locus results to multiple loci. It would be of considerable interest to find out for which genetic architectures and migration patterns such constraints occur.

It remains an unresolved problem how many alleles at how many loci can be maintained segregating under the assumption of DIDID or under the stronger assumption of no $G \times E$ interaction. Theorem 3.14 and Corollary 3.9 in N09b imply that if there is DIDID and the number of alleles is the same at each locus, then Γ is a generic upper bound on the number of segregating alleles per locus.

Eventually, one would like to quantify how frequent polymorphism is in the Levene model without epistasis, and how this depends on the selection scheme and the dominance relations. Numerical results for two diallelic loci and two demes show that for arbitrary intermediate dominance, the volume of the parameter space in which a full polymorphism is maintained varies between about 12% (for very weak selection) and about 16% (for very strong selection). Under the assumption of sublinear fitnesses (see (3.5) and Corollary 3.3(c)), these numbers increase to 22% and 36%, respectively. In the absence of $G \times E$ interaction, they are 0% and

43%, respectively (Bürger, 2009c). With increasing number of alleles or loci, the volume of parameter space in which a stable internal equilibrium exists will certainly decrease rapidly, at least if the parameter range is unconstrained. Still, the set of parameter combinations for which a significant fraction of loci can be maintained polymorphic may be quite large.

Therefore, we expect that spatially varying selection has the potential to maintain considerable multilocus polymorphism, especially if locally beneficial alleles tend to be partially dominant. Properties of the genetic architecture, such as the dominance pattern or the presence or absence of (certain forms of) $G \times E$ interaction, may greatly constrain this potential. In view of this, it seems worthwhile to study phenotypic plasticity and the evolution of genetic architecture in a spatial context (see e.g. Zhivotovsky et al., 1996; de Jong, 1999; Otto and Bourguet, 1999; de Jong and Gavrilets, 2000; van Doorn and Dieckmann, 2006).

Whereas DIDID plays an important role in the Levene model in constraining the amount of polymorphism, this is not necessarily so for general migration patterns. Peischl (2010) studied a single-locus model with two demes and general migration. He proved that, if there is DIDID, three alleles can be maintained at a stable equilibrium. His numerical computations suggest that more than three alleles can be maintained for an open set of parameters. His results have interesting interpretations in the context of the coexistence of specialists and generalists. Further, it is yet unknown whether DIDID also plays a prominent role in the Levene model with epistasis, for instance, if there is spatially varying stabilizing selection on a quantitative trait. This reconfirms the conviction that studying the role of genetic architecture for the maintenance of polymorphism in a spatial context seems to be a worthwhile enterprise.

Acknowledgments

The author is most grateful to Professor Thomas Nagylaki for extensive conversation and sharing of unpublished results. Many thanks go to Tom Nagylaki, Nick Barton, and an anonymous reviewer for perceptive comments on previous versions of this paper. This work was supported by a grant of the Vienna Science and Technology Fund (WWTF) and by grant P21305 of the Austrian Science Fund (FWF).

Appendix A. Proof of Eq. (3.7)

Because we assume DIDID (3.6), (2.8) yields

$$\xi^{(n)}_{i_nj_n} = \vartheta^{(n)}_{i_nj_n}\xi^{(n)}_{i_ni_n} + \vartheta^{(n)}_{j_ni_n}\xi^{(n)}_{j_nj_n} \tag{A.1}$$

and

$$\xi^{(n)}_{i_n} = \xi^{(n)}_{i_ni_n}\sum_{j_n}\vartheta^{(n)}_{i_nj_n}p^{(n)}_{j_n} + \sum_{j_n}\vartheta^{(n)}_{j_ni_n}p^{(n)}_{j_n}\xi^{(n)}_{j_nj_n}. \tag{A.2}$$

For every locus n, we introduce the nonnegative $I_n \times I_n$ matrices $M^{(n)}$, $D^{(n)}$, $P^{(n)}$, and $S^{(n)}$ as follows:

$$M^{(n)} = D^{(n)}P^{(n)} + S^{(n)}, \tag{A.3}$$

where

$$[D^{(n)}]_{i_nj_n} = \vartheta^{(n)}_{j_ni_n}, \tag{A.4}$$

$P^{(n)}$ is diagonal with

$$[P^{(n)}]_{i_ni_n} = p^{(n)}_{i_n}, \tag{A.5}$$

and $S^{(n)}$ is diagonal with

$$[S^{(n)}]_{i_ni_n} = \sum_{j_n}\vartheta^{(n)}_{i_nj_n}p^{(n)}_{j_n}. \tag{A.6}$$

A simple calculation shows that $M^{(n)}$ is stochastic:

$$\begin{aligned}\sum_{j_n}[M^{(n)}]_{i_nj_n} &= \left(\sum_{j_n}\vartheta^{(n)}_{j_ni_n}p^{(n)}_{j_n} + \sum_{j_n}\vartheta^{(n)}_{i_nj_n}p^{(n)}_{j_n}\right)\\ &= \sum_{j_n}(\vartheta^{(n)}_{j_ni_n} + \vartheta^{(n)}_{i_nj_n})p^{(n)}_{j_n} = 1.\end{aligned} \tag{A.7}$$

Because we consider only internal equilibria, we have $p^{(n)}_{j_n} > 0$ for every h and j_n. Hence, $P^{(n)}$ is nonsingular. Furthermore,

$$[S^{(n)}]_{i_ni_n} \geq \vartheta^{(n)}_{i_ni_n}p^{(n)}_{i_n} = \frac{1}{2}p^{(n)}_{i_n} > 0, \tag{A.8}$$

whence $S^{(n)}$ is also nonsingular.

Lemma A.1. *If $\hat\rho$ is a gene-frequency equilibrium and $M^{(n)} = M^{(n)}(\hat\rho)$ is nonsingular, then*

$$\hat\xi^{(n)}_{i_nj_n} = \hat{\bar\xi}^{(n)} \quad \textit{holds for every } i_n \textit{ and } j_n. \tag{A.9}$$

Proof. Let $x^{(n)} = (\hat\xi^{(n)}_{11}, \ldots, \hat\xi^{(n)}_{I_nI_n})^T$. Then, in vector form, (A.2) reads

$$\hat\xi^{(n)}_{i_n} = [S^{(n)}x^{(n)}]_{i_n} + [D^{(n)}P^{(n)}x^{(n)}]_{i_n} = [M^{(n)}x^{(n)}]_{i_n}, \tag{A.10}$$

where $\hat\xi^{(n)}_{i_n}$, $x^{(n)}$, and $M^{(n)}$ all depend on $\hat\rho$. Since the internal gene-frequency equilibria $\hat\rho$ are precisely the solutions of (2.15), they are precisely the solutions of

$$M^{(n)}x^{(n)} = \hat{\bar\xi}^{(n)}\mathbf{1}^{(n)}, \tag{A.11}$$

where $\mathbf{1}^{(n)} = (1, \ldots, 1)^T \in \mathbb{R}^{I_n}$. Because $M^{(n)}$ is stochastic, $x^{(n)} = \hat{\bar\xi}^{(n)}\mathbf{1}^{(n)}$ is a solution of (A.11). If $M^{(n)}$ is nonsingular, for any given $\bar\xi^{(n)}$, it is the unique solution. Therefore, $\hat\xi^{(n)}_{i_ni_n} = \hat{\bar\xi}^{(n)}$ is satisfied for every i_n. The assertion now follows from (A.1) and (3.6b). □

Our aim now is to show that $M^{(n)}$ is nonsingular. We write U_n for the $n \times n$ matrix with all entries 1 and recall that a real matrix B is skew-symmetric if $B + B^T = 0$.

Lemma A.2. *Let B_n be a skew-symmetric $n \times n$ matrix and $C_n(x) = B_n + xU_n$. Then $\det(C_n(x)) \geq 0$ for every $n \geq 1$ and $x \geq 0$.*

Proof. We begin by deriving a recursive formula for $\det(C_n(x))$. If we first subtract row n from all other rows of $C_n(x)$ and then (the resulting) column n from all other columns, we obtain the matrix

$$\begin{pmatrix} & & & b_{1n}\\ & \tilde B_{n-1} & & \vdots\\ & & & b_{n-1,n}\\ b_{n1} & b_{n2} & \ldots & b_{n,n-1} & x\end{pmatrix}, \tag{A.12}$$

where the entry (k, j) of the matrix $\tilde B_{n-1}$ is

$$\tilde b_{kj} = b_{kj} - b_{nj} - b_{kn}. \tag{A.13}$$

Because $B_n + B_n^T = 0$, a simple calculation yields $\tilde B_{n-1} + \tilde B_{n-1}^T = 0$. Clearly, if $x = 0$, the determinant of the matrix (A.12) is simply $\det(B_n)$. Developing the determinant of (A.12) with respect to the last row (or column), it follows immediately that

$$\det(C_n(x)) = x\det(\tilde B_{n-1}) + \det(B_n). \tag{A.14}$$

Now we use the fact (Godsil, 1993, pp. 113–116) that for every skew-symmetric matrix B_n,

$$\det(B_{2n-1}) = 0, \tag{A.15a}$$

$$\det(B_{2n}) = (\mathrm{Pf}(B_{2n}))^2 \tag{A.15b}$$

holds, where

$$\mathrm{Pf}(B_{2n}) = \frac{1}{2^n n!}\sum_{\sigma\in S_{2n}} \mathrm{sgn}(\sigma)\prod_{i=1}^{n} b_{\sigma(2i-1),\sigma(2i)} \tag{A.16}$$

denotes the Pfaffian of the matrix B_{2n}, and S_{2n} is the permutation group of a set of $2n$ elements. Applying (A.15) to (A.14), we obtain

$$\det(C_n(x)) = \begin{cases} (\mathrm{Pf}(B_n))^2 & \text{if } n \text{ is even,} \\ x(\mathrm{Pf}(\tilde{B}_{n-1}))^2 & \text{if } n \text{ is odd,} \end{cases} \tag{A.17}$$

which proves the lemma. □

Lemma A.3. *For arbitrary number I_n of alleles, arbitrary matrix $D^{(n)}$ of dominance parameters satisfying (3.6b), and arbitrary positive matrix $P^{(n)}$, the matrix $M^{(n)}$ is nonsingular.*

Proof. We shall prove $\det(M^{(n)}) > 0$. For the proof, we omit the locus superscripts (n) and assume that all matrices are $n \times n$. Because S is nonsingular, (A.3)–(A.7) show that

$$M = (DPS^{-1} + I)S, \tag{A.18}$$

where I is the $n \times n$ identity matrix. By assumption (3.6b), D can be written as $D = B + \frac{1}{2}U$, where B is skew-symmetric. Therefore, Lemma A.2 shows that $\det(D) \geq 0$. Because this holds for arbitrary dimension, every principal minor (of arbitrary order) of D is also nonnegative. Since PS^{-1} is a diagonal matrix with positive elements on the diagonal, we have $\det(DPS^{-1}) = \det(D)\det(PS^{-1}) \geq 0$. Because this holds for every dimension and because PS^{-1} is diagonal, the principal minors of DPS^{-1} are the products of the corresponding principal minors of D and PS^{-1} (the latter being just products of the corresponding diagonal elements). Hence, all principal minors of DPS^{-1} are nonnegative.

It is another well-known fact (e.g. Gantmacher, 1986, p. 99) that the coefficient of x^k of the characteristic polynomial of a square matrix can be written as the sum of all principal minors of order $n - k$. As a consequence, $\det(I + DPS^{-1})$, which equals the characteristic polynomial of DPS^{-1} evaluated at -1, can be written as 1 plus the sum of all principal minors of order ≥ 1. Since by the above reasoning all of them are nonnegative, we have

$$\det(M) = \det(DPS^{-1} + I)\det(S) \geq \det(S) > 0, \tag{A.19}$$

which finishes the proof. □

Lemmas A.1 and A.3 show that Eq. (3.7) holds for every internal gene-frequency equilibrium if there is DIDID.

Appendix B. Proof of Proposition 5.4

Throughout, we assume two alleles per locus and DIDID, i.e., (5.7). Let $\mathsf{M}_{L,\Gamma}$ denote the set of all $L \times \Gamma$ matrices with strictly positive entries. For given alleles i, j at locus n, we denote by $U_{ij} = (u^{(n)}_{ij,\alpha}) \in \mathsf{M}_{L,\Gamma}$ the matrix containing all fitness contributions $u^{(n)}_{ij,\alpha}$, where $u^{(n)}_{ij,\alpha} > 0$ and $n \in \mathsf{L}$ and $\alpha \in \mathsf{G}$. Further, let

$$\mathsf{G}' = \{1, \ldots, \Gamma - 1\} \tag{B.1}$$

and denote

$$c \in \mathsf{C} = \left\{ (c_1, \ldots, c_{\Gamma-1})^T : c_\alpha > 0, \alpha \in \mathsf{G}', \text{ and } \sum_{\alpha\in\mathsf{G}'} c_\alpha < 1 \right\}. \tag{B.2}$$

Then, $c_\Gamma = 1 - \sum_{\alpha\in\mathsf{G}'} c_\alpha > 0$ and $(c_1, \ldots, c_\Gamma)^T \in \mathrm{int}\,\Delta_\Gamma$.

For any given matrix $A = (a_1, \ldots, a_\Gamma) \in \mathsf{M}_{L,\Gamma}$, where $a_\alpha \in \mathbb{R}^L$ $(\alpha \in \mathsf{G})$ is a column vector, we write

$$A' = (a_1, \ldots, a_{\Gamma-1}) \in \mathsf{M}_{L,\Gamma-1}$$

for the matrix obtained by omitting the last column (which requires $\Gamma \geq 2$). For any given vector a, $a > 0$ means that every component of a is positive, by which we always mean strictly positive.

We shall need the following parameter sets:

$$\mathsf{P}_\mathsf{D} = \{(c, U_{11}, U_{22}) : c \in \mathsf{C},\ U_{11}, U_{22} \in \mathsf{M}^+_{L,\Gamma}\}, \tag{B.3a}$$

$$\mathsf{P}'_\mathsf{D} = \{(c, U'_{11}, U_{22}) : c \in \mathsf{C},\ U'_{11} \in \mathsf{M}^+_{L,\Gamma-1},\ U_{22} \in \mathsf{M}^+_{L,\Gamma}\}. \tag{B.3b}$$

Because we assume (5.7), the gene-frequency dynamics (5.5a) is completely specified by $\vartheta \in [0,1]^L$ and $(c, U_{11}, U_{22}) \in \mathsf{P}_\mathsf{D}$.

Our aim is to prove Proposition 5.4, which here we formulate more precisely.

Proposition B.1. *If $1 \leq L \leq \Gamma - 1$ and $\vartheta \in [0,1]^L$. Then there is an open, nonempty subset $\mathsf{U} \subset \mathsf{P}_\mathsf{D}$, such that for every $(c, U_{11}, U_{22}) \in \mathsf{U}$ an isolated, internal, asymptotically stable equilibrium point of the gene-frequency dynamics (5.5a) exists. It is the unique solution $\hat{\rho}$ of (5.8) in $\mathrm{int}\,\Omega = (0,1)^L$.*

The proof is based on three lemmas.

Lemma B.2. *$\hat{\rho} = \frac{\mathbf{1}}{2} = (\frac{1}{2}, \ldots, \frac{1}{2})^T$ is an equilibrium of the gene-frequency dynamics (5.5a) if and only if*

$$\sum_\alpha c_\alpha \frac{u^{(n)}_{11,\alpha} - u^{(n)}_{22,\alpha}}{w^*_\alpha} = 0 \quad \textit{for every } n \in \mathsf{L}, \tag{B.4}$$

where

$$w^*_\alpha = \overline{w}_\alpha\left(\tfrac{\mathbf{1}}{2}\right) = \frac{1}{4}\sum_n \big[(1 + 2\vartheta^{(n)})u^{(n)}_{11,\alpha} + (3 - 2\vartheta^{(n)})u^{(n)}_{22,\alpha}\big]. \tag{B.5}$$

Proof. Because we assume (5.7), the internal equilibria of (5.5a) are precisely the solutions of (5.8). Substitution of $\hat{\rho} = \frac{\mathbf{1}}{2}$ and (5.7) into (5.4) and (2.5) yields (B.5). □

We denote $u_{11,\Gamma} = (u^{(1)}_{11,\Gamma}, \ldots, u^{(L)}_{11,\Gamma})^T \in \mathbb{R}^L$.

Lemma B.3. *Let $\vartheta \in [0,1]^L$. There exists an open subset $\mathsf{O}' \subset \mathsf{P}'_\mathsf{D}$ such that for every $(c, U'_{11}, U_{22}) \in \mathsf{O}'$ a unique solution $u_{11,\Gamma}$ of (B.4) exists. It satisfies $u_{11,\Gamma} > 0$.*

Proof. Let

$$\lambda_n = -\frac{1}{c_\Gamma}\sum_{\alpha\in\mathsf{G}'} \frac{c_\alpha}{w^*_\alpha}(u^{(n)}_{11,\alpha} - u^{(n)}_{22,\alpha}), \tag{B.6}$$

$x_n = u^{(n)}_{11,\Gamma}$, $x = \sum_n (1 + 2\vartheta^{(n)})x_n$, and $\mu_1 = \sum_n (3 - 2\vartheta^{(n)})u^{(n)}_{22,\Gamma}$. Then we obtain $w^*_\Gamma = \frac{1}{4}(x + \mu_1)$ from (B.4), and (B.5) yields

$$\frac{c_\Gamma}{w^*_\Gamma}(x_n - u^{(n)}_{22,\Gamma}) = c_\Gamma \lambda_n. \tag{B.7}$$

Hence,

$$x_n = u^{(n)}_{22,\Gamma} + \frac{1}{4}\lambda_n(x + \mu_1), \quad n \in \mathsf{L}. \tag{B.8}$$

We solve this system for (x_n). Let $\lambda = \sum_n (1 + 2\vartheta^{(n)})\lambda_n$ and $\mu_2 = \sum_n (1 + 2\vartheta^{(n)})u^{(n)}_{22,\Gamma}$.

Multiplying (B.8) by $(1 + 2\vartheta^{(n)})$ and summing over all n, we obtain

$$x = \mu_2 + \frac{1}{4}\lambda(x + \mu_1). \tag{B.9}$$

If $\lambda \neq 4$, (B.9) has the unique solution

$$x = \frac{\lambda\mu_1 + 4\mu_2}{4 - \lambda}. \tag{B.10}$$

Substituting (B.10) into (B.8), we obtain

$$x_n = u^{(n)}_{22,\Gamma} + \frac{\lambda_n}{4-\lambda}(\mu_1+\mu_2) = u^{(n)}_{22,\Gamma} + \frac{4\lambda_n}{4-\lambda}\sum_l u^{(l)}_{22,\Gamma}, \quad \text{(B.11)}$$

which exists and is uniquely determined if $\lambda \neq 4$. This yields the desired solution $u_{11,\Gamma} = (x_1, \ldots, x_n)^T$. It satisfies $u_{11,\Gamma} > 0$ if

$$\left|\frac{4\lambda_n}{4-\lambda}\right| \sum_l u^{(l)}_{22,\Gamma} < u^{(n)}_{22,\Gamma} \quad \text{for every } n \in \mathsf{L}. \quad \text{(B.12)}$$

It remains to show that (B.12) holds on an open set of parameters. Because $|\lambda| \leq 3\sum_n |\lambda_n|$, we have $\lambda_n/(4-\lambda) \leq 1$ whenever $\sum_n |\lambda_n| \leq 1$. The latter is easy to achieve by choosing the c_α, $\alpha \in \mathsf{G}'$, sufficiently small. As a consequence and because $u^{(n)}_{22,\Gamma} > 0$ for every n, the estimate (B.12) can be realized for every ϑ and every (U'_{11}, U_{22}) by choosing a sufficiently small $\epsilon > 0$ such that $\max_{\alpha\in\mathsf{G}'} c_\alpha < \epsilon$. Choosing (U'_{11}, U_{22}) from an arbitrary, but bounded, open subset $\mathsf{M} \subset \mathsf{M}^+_{L,\Gamma-1} \times \mathsf{M}^+_{L,\Gamma}$, we can find a uniformly small $\epsilon > 0$. Thus, we can choose the desired set $\mathsf{O}' = \mathsf{C}_\epsilon \times \mathsf{M}$, where $\mathsf{C}_\epsilon = (0,\epsilon)^{\Gamma-1}$ for sufficiently small $\epsilon > 0$. Thus, for every $\vartheta \in [0,1]^L$, there exists an open bounded set O' and a vector $u_{11,\Gamma}$ with the desired properties. □

Remark B.4. $u_{11,\Gamma}$ depends smoothly on (c, U'_{11}, U_{22}) and on ϑ.

Remark B.5. The set O' can be chosen such that it applies to every $\vartheta \in [0,1]^L$. This follows from the proof because $\frac{1}{4}\sum_n(u^{(n)}_{11,\alpha} + u^{(n)}_{22,\alpha}) \leq w^*_\alpha \leq \frac{3}{4}\sum_n(u^{(n)}_{11,\alpha} + u^{(n)}_{22,\alpha})$ and, therefore, the critical ϵ can be chosen independently of ϑ. Hence, the left-hand side in (B.12) can be made small uniformly in $\vartheta \in [0,1]^L$.

Next we show that small perturbations of $u_{11,\Gamma}$ yield equilibria $\hat\rho$ close to $\frac{1}{2}$. For this, we need the Hessian matrix of $F = \ln\tilde w$, evaluated at $\rho \in \operatorname{int}\Omega$. We denote it by $H(\rho) \in \mathsf{M}_{L,L}$, i.e.,

$$H_{nl}(\rho) = \frac{\partial^2 F(\rho)}{\partial p^{(n)}\,\partial p^{(l)}}. \quad \text{(B.13)}$$

A critical point ρ of F is isolated if $H(\rho)$ is invertible. Because we need to check regularity of H at critical points, we calculate $H(\rho)$. Partial differentiation of (2.14), using (2.5) and (5.4), yields

$$\frac{\partial F}{\partial p^{(n)}} = 2\sum_\alpha \frac{c_\alpha}{\overline{w}_\alpha}(u^{(n)}_{1,\alpha} - u^{(n)}_{2,\alpha}). \quad \text{(B.14)}$$

Differentiating this with respect to $p^{(l)}$, we obtain straightforwardly from (5.3) and (5.4):

$$\begin{aligned} H_{nl}(\rho) &= 2\sum_\alpha c_\alpha \frac{\partial}{\partial p^{(l)}}\left(\frac{u^{(n)}_{1,\alpha} - u^{(n)}_{2,\alpha}}{\overline{w}_\alpha}\right) \\ &= 2\sum_\alpha c_\alpha \left[\delta_{nl}\frac{u^{(n)}_{11,\alpha} - 2u^{(n)}_{12,\alpha} + u^{(n)}_{22,\alpha}}{\overline{w}_\alpha}\right. \\ &\quad \left. - 2\frac{(u^{(n)}_{1,\alpha} - u^{(n)}_{2,\alpha})(u^{(l)}_{1,\alpha} - u^{(l)}_{2,\alpha})}{\overline{w}^2_\alpha}\right], \end{aligned} \quad \text{(B.15)}$$

where $\delta_{nl} = 1$ if $n = l$, and $\delta_{nl} = 0$ otherwise.

Now we invoke (5.7) and evaluate $H_{nl}(\rho)$ at $\rho = \frac{1}{2}$. This gives $u^{(n)}_{11,\alpha} - 2u^{(n)}_{12,\alpha} + u^{(n)}_{22,\alpha} = (1-2\vartheta^{(n)})(u^{(n)}_{11,\alpha} - u^{(n)}_{22,\alpha})$ and $\overline{w}_\alpha = w^*_\alpha$. Therefore, (B.4) implies

$$\sum_\alpha c_\alpha \frac{u^{(n)}_{11,\alpha} - 2u^{(n)}_{12,\alpha} + u^{(n)}_{22,\alpha}}{w^*_\alpha} = 0. \quad \text{(B.16)}$$

Applying (5.9) with $\rho = \frac{1}{2}$ to the second term in (B.15), the entries of the Hessian matrix evaluated at $\frac{1}{2}$, $\Phi = H(\frac{1}{2})$, become

$$\Phi_{nl} = -\sum_\alpha c_\alpha \frac{(u^{(n)}_{11,\alpha} - u^{(n)}_{22,\alpha})(u^{(l)}_{11,\alpha} - u^{(l)}_{22,\alpha})}{(w^*_\alpha)^2}. \quad \text{(B.17)}$$

Now we have all ingredients to formulate and prove our last lemma.

Lemma B.6. *Let $1 \leq L \leq \Gamma - 1$ and $\vartheta \in [0,1]^L$. Then there exists an open subset $\mathsf{O}'_1 \subset \mathsf{O}'$ such that for every $(c, U'_{11}, U_{22}) \in \mathsf{O}'_1$ and corresponding solution $u_{11,\Gamma} > 0$ of (B.4), the Hessian Φ is invertible.*

Proof. Let C denote the $\Gamma \times \Gamma$ diagonal matrix with entries c_α along the diagonal, and $C^{\frac{1}{2}}$ its square root, which exists and is invertible because $c_\alpha > 0$ for every α. Moreover, we set

$$e^{(n)}_\alpha = \frac{u^{(n)}_{11,\alpha} - u^{(n)}_{22,\alpha}}{w^*_\alpha} \quad \text{(B.18)}$$

and define E as the matrix with these entries, i.e., $E = (e^{(n)}_\alpha) \in \mathsf{M}_{L,\Gamma}$. Then we can write

$$\Phi = -ECE^T = -(EC^{1/2})(EC^{1/2})^T. \quad \text{(B.19)}$$

Because (B.4) informs us that $\sum_\alpha c_\alpha e^{(n)}_\alpha = 0$ for every n, generically, E has rank $\min(\Gamma-1, L)$. If $\Gamma \geq L+1$, then, generically, rank $E = L$. Because rank $MM^T = \operatorname{rank} M$ is valid for every matrix $M \in \mathsf{M}_{L,\Gamma}$, we obtain rank $\Phi = \operatorname{rank}(C^{1/2}E) = \operatorname{rank} E$, where the latter equality holds because multiplication by a nonsingular matrix leaves the rank unchanged. Therefore, Φ is (generically) invertible if and only if $\Gamma \geq L+1$.

Finally, if Q' denotes the preimage of the open subset of $\mathsf{M}_{L,\Gamma}$ of matrices of rank L under the continuous map $(U'_{11}, U_{22}) \mapsto E$, then we can choose $\mathsf{O}'_1 = (\mathsf{C}_\epsilon \times \mathsf{Q}') \cap \mathsf{O}'$, where O' is as in the proof of Lemma B.3. Therefore the set O'_1 is open and nonempty. □

Remark B.7. In the absence of DIDID, this Hessian matrix is of the form $\Phi = D - E^T CE$, where D is a nonnegative diagonal matrix that vanishes if there is DIDID; cf. (B.15). Then Φ is generically invertible whenever $L \geq 1$ and $\Gamma \geq 2$. In fact, the proof can be extended to arbitrary dominance relations, thus giving a very different proof of Result 5.1 that does not assume weak selection. However, the generalizations of Lemmas B.3 and B.6 require considerable work.

Proof of Proposition B.1. We choose (c, U'_{11}, U_{22}) and the corresponding $u_{11,\Gamma}$ according to Lemma B.6, i.e., $(c, U'_{11}, U_{22}) \in \mathsf{O}'_1$. Clearly, $(c, U_{11}, U_{22}) \in \mathsf{P}_\mathsf{D}$. By Lemma B.2, $\frac{1}{2}$ is an internal equilibrium of (5.5a), and by Lemma B.6 the Hessian $H(\frac{1}{2}) = \Phi$ is invertible. Therefore, $\frac{1}{2}$ is an isolated equilibrium of (5.5a). Since (B.15), together with (5.3), (5.4), (2.5) and Remark B.4, demonstrates that $H(\rho)$ is smooth in ρ, the Implicit Function Theorem shows that there exists an open neighborhood $\mathsf{U} \subset \mathsf{P}_\mathsf{D}$ of (c, U_{11}, U_{22}) for which there is a unique, isolated gene-frequency equilibrium $\hat\rho$ close to $\frac{1}{2}$. Theorem 3.14 in N09b yields global asymptotic stability. □

References

Akin, E., 1993. The General Topology of Dynamical Systems. Amer. Math. Soc, Providence, RI.

Barton, N.H., 2010. What role does natural selection play in speciation? Phil. Trans. R. Soc. B. 365, 1825–1840.

Bürger, R., 2000. The Mathematical Theory of Selection, Recombination, and Mutation. Wiley, Chichester.

Bürger, R., 2009a. Multilocus selection in subdivided populations I. Convergence properties for weak or strong migration. J. Math. Biol. 58, 939–978.

Bürger, R., 2009b. Multilocus selection in subdivided populations II. Maintenance of polymorphism for weak or strong migration. J. Math. Biol. 58, 979–997.

Bürger, R., 2009c. Polymorphism in the two-locus Levene model with nonepistatic directional selection. Theor. Popul. Biol. 76, 214–228.

Christiansen, F.B., Feldman, M., 1975. Subdivided populations: a review of the one- and two-locus deterministic theory. Theor. Popul. Biol. 7, 13–38.

Christiansen, F.B., 1999. Population Genetics of Multiple Loci. Wiley, Chichester.

Conley, C., 1978. Isolated invariant sets and the Morse index. In: NSF CBMS Lecture Notes, vol. 38. Amer. Math. Soc, Providence, RI.

de Jong, G., 1999. Unpredictable selection in a structured population leads to local differentiation in evolved reaction norms. J. Evol. Biol. 12, 839–851.

de Jong, G., Gavrilets, S., 2000. Maintenance of genetic variation in phenotypic plasticity: the role of environmental variation. Genet. Res. 76, 295–304.

Ewens, W., 2004. Mathematical Population Genetics. I. Theoretical Introduction, second ed.. Springer, Berlin.
Gantmacher, F.R., 1986. Matrizentheorie. Springer, Berlin.
Geiringer, H., 1944. On the probability theory of linkage in Mendelian heredity. Ann. Math. Stat. 15, 25–57.
Godsil, C.D., 1993. Algebraic Combinatorics. Chapman and Hall, New York.
Karlin, S., 1977. Gene frequency patterns in the Levene subdivided population model. Theor. Popul. Biol. 11, 356–385.
Karlin, S., 1978. Theoretical aspects of multilocus selection balance I. In: Studies in Mathematics. In: Levin, S.A. (Ed.), Studies in Mathematical Biology, vol. 16. The Mathematical Association of America, Washington, pp. 503–587.
Karlin, S., 1982. Classification of selection-migration structures and conditions for a protected polymorphism. Evol. Biol. 14, 61–204.
Karlin, S., Liberman, U., 1979. Representation of nonepistatic selection models and analysis of multilocus Hardy-Weinberg equilibrium configurations. J. Math. Biol. 7, 353–374.
Karlin, S., McGregor, J., 1972. Polymorphism for genetic and ecological systems with weak coupling. Theor. Popul. Biol. 3, 210–238.
Kirkpatrick, M., Johnson, T., Barton, N., 2002. General models of multilocus selection. Genetics 161, 1727–1750.
Kirzhner, V., Lyubich, Yu.I., 1997. Multilocus dynamics under haploid selection. J. Math. Biol. 35, 391–408.
Kun, L.A., Lyubich, Yu.I., 1979. Convergence to equilibrium under the action of additive selection in a multilocus multiallelic population. Sov. Math. Dokl. 20, 1380–1382.
Kun, L.A., Lyubich, Yu.I., 1980. Convergence to equilibrium in a polylocus polyallele population with additive selection. Probl. Inform. Transmiss. 16, 152–161.
LaSalle, J.P., 1977. Stability theory for difference equations. In: Hale, J.K. (Ed.), Studies in Ordinary Differential Equations. In: Studies in Mathematics, vol. 14. Mathematical Association of America, Washington, DC, pp. 1–31.
Levene, H., 1953. Genetic equilibrium when more than one ecological niche is available. Amer. Natur. 87, 331–333.
Li, W.-H., Nei, M., 1974. Stable linkage disequilibrium without epistasis in subdivided populations. Theor. Popul. Biol. 6, 173–183.
Lyubich, Yu.I., 1992. Mathematical Structures in Population Genetics. Springer, Berlin, Heidelberg, New York.
Nagylaki, T., 1992. Introduction to Theoretical Population Genetics. Springer, Berlin, Heidelberg, New York.
Nagylaki, T., 1993. The evolution of multilocus systems under weak selection. Genetics 134, 627–647.
Nagylaki, T., 2009a. Polymorphism in multiallelic migration–selection models with dominance. Theor. Popul. Biol. 75, 239–259.
Nagylaki, T., 2009b. Evolution under the multilocus Levene model. Theor. Popul. Biol. 76, 197–213.
Nagylaki, T., Hofbauer, J., Brunovský, P., 1999. Convergence of multilocus systems under weak epistasis or weak selection. J. Math. Biol. 38, 103–133.
Nagylaki, T., Lou, Y., 2008. The dynamics of migration–selection models. In: Friedman, A. (Ed.), Tutorials in Mathematical Biosciences IV. In: Lect. Notes Math., vol. 1922. Springer, Berlin, Heidelberg, New York, pp. 119–172.
Peischl, S., 2010. Dominance and the maintenance of polymorphism in multiallelic migration–selection models with two demes. Theor. Popul. Biol. 78, 12–25.
Roze, D., Rousset, F., 2008. Multilocus models in the infinite island model of population structure. Theor. Popul. Biol. 73, 529–542.
Spichtig, M., Kawecki, T.J., 2004. The maintenance (or not) of polygenic variation by soft selection in a heterogeneous environment. Amer. Natur. 164, 70–84.
Strobeck, C., 1979. Haploid selection with n alleles and m niches. Am. Nat. 113, 439–444.
Turelli, M., Barton, N.H., 1990. Dynamics of polygenic characters under selection. Theor. Popul. Biol. 38, 1–57.
van Doorn, G.S., Dieckmann, U., 2006. The long-term evolution of multilocus traits under frequency-dependent disruptive selection. Evolution 60, 2226–2238.
Wiehe, T., Slatkin, M., 1998. Epistatic selection in a multi-locus Levene model and implications for linkage disequilibrium. Theor. Popul. Biol. 53, 75–84.
Zhivotovsky, L.A., Feldman, M.W., Bergman, A., 1996. On the evolution of phenotypic plasticity in a spatially heterogeneous environment. Evolution 50, 547–558.
Zhivotovsky, L.A., Gavrilets, S., 1992. Quantitative variability and multilocus polymorphism under epistatic selection. Theor. Popul. Biol. 42, 254–283.